

Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception

Charles Chubb and George Sperling

Human Information Processing Laboratory, Psychology Department, New York University, 6 Washington Place,
New York, New York 10003

Received November 17, 1987; accepted June 7, 1988

To some degree, all current models of visual motion-perception mechanisms depend on the power of the visual signal in various spatiotemporal-frequency bands. Here we show how to construct counterexamples: visual stimuli that are consistently perceived as obviously moving in a fixed direction yet for which Fourier-domain power analysis yields no systematic motion components in any given direction. We provide a general theoretical framework for investigating non-Fourier motion-perception mechanisms; central are the concepts of drift-balanced and microbalanced random stimuli. A random stimulus S is drift balanced if its expected power in the frequency domain is symmetric with respect to temporal frequency, that is, if the expected power in S of every drifting sinusoidal component is equal to the expected power of the sinusoid of the same spatial frequency, drifting at the same rate in the opposite direction. Additionally, S is microbalanced if the result WS of windowing S by any space-time-separable function W is drift balanced. We prove that (i) any space-time-separable random (or nonrandom) stimulus is microbalanced; (ii) any linear combination of pairwise independent microbalanced (respectively, drift-balanced) random stimuli is microbalanced and drift balanced if the expectation of each component is uniformly zero; (iii) the convolution of independent microbalanced and drift-balanced random stimuli is microbalanced and drift balanced; (iv) the product of independent microbalanced random stimuli is microbalanced; and (v) the expected response of any Reichardt detector to any microbalanced random stimulus is zero at every instant in time. Examples are provided of classes of microbalanced random stimuli that display consistent and compelling motion in one direction. All the results and examples from the domain of motion perception are transposable to the space-domain problem of detecting orientation in a texture pattern.

1. INTRODUCTION

Central to the study of human visual motion perception is the relationship between perceived motion and the Fourier transform of the spatiotemporal visual stimulus. Points in the domain of the spatiotemporal Fourier transform correspond to drifting sinusoidal gratings. For a wide range of spatial and temporal frequencies, such drifting sinusoids are perceived to move uniformly across the visual field, and their apparent speed and direction are direct functions of spatiotemporal frequency. For the most part, the motion displayed by simple linear combinations of such gratings reflects quite reasonably the individual contributions of the components.^{1,2}

Indeed, current models of human motion perception implicitly or explicitly involve some degree of Fourier decomposition (bandpass filtering) of the image stream.¹⁻⁶ Generally, of course, the decomposition is localized to finite temporal intervals and subregions of the visual field.

It has long been realized, however, that certain sorts of apparent motion cannot be understood directly in terms of their power spectra.⁷⁻¹⁴ For instance, much attention has been focused on sums of drifting gratings of slightly different, high spatial frequencies.¹⁰⁻¹² In general, the perceived velocity of such stimuli is determined not directly by the frequencies of the summed components but by the pattern of beats at their difference frequency.

Sperling¹³ demonstrated "movement without correlation" in a different stimulus whose Fourier transform, when com-

puted globally or locally, contained no consistent moving components and yet was perceived to move decisively in a fixed direction. Subsequently, Petersik *et al.*¹⁴ studied similar displays in an effort to clarify the relationship between stage 1 (autocorrelational, Fourier) mechanisms and the higher-order stage 2 mechanisms mediating the perception of what we call¹⁵ non-Fourier motion.

The purpose of this paper is to provide (i) a general theoretical basis and (ii) an array of specific tools for studying non-Fourier motion-perception mechanisms.¹⁶

2. ANALYZING A STIMULUS: INTUITIVE FOURIER DECOMPOSITION

We begin with a brief, informal discussion to show how particular motion stimuli can be analyzed into drifting sinusoids. For illustration we use one-dimensional spatiotemporal stimuli that move either to the left or to the right and whose luminance varies in only the horizontal dimension, although all the results that we derive apply in all cases to stimuli of two spatial dimensions and time. A one-dimensional, horizontally moving stimulus is represented conveniently by a two-dimensional function $l(x, t)$, where x (the horizontal axis) indicates the spatial pattern of luminance and t (the vertical axis, with time increasing upward) indicates the temporal luminance pattern. In this representation, usually it is immediately obvious which way the dominant Fourier components of l tend to slope (up and to the left

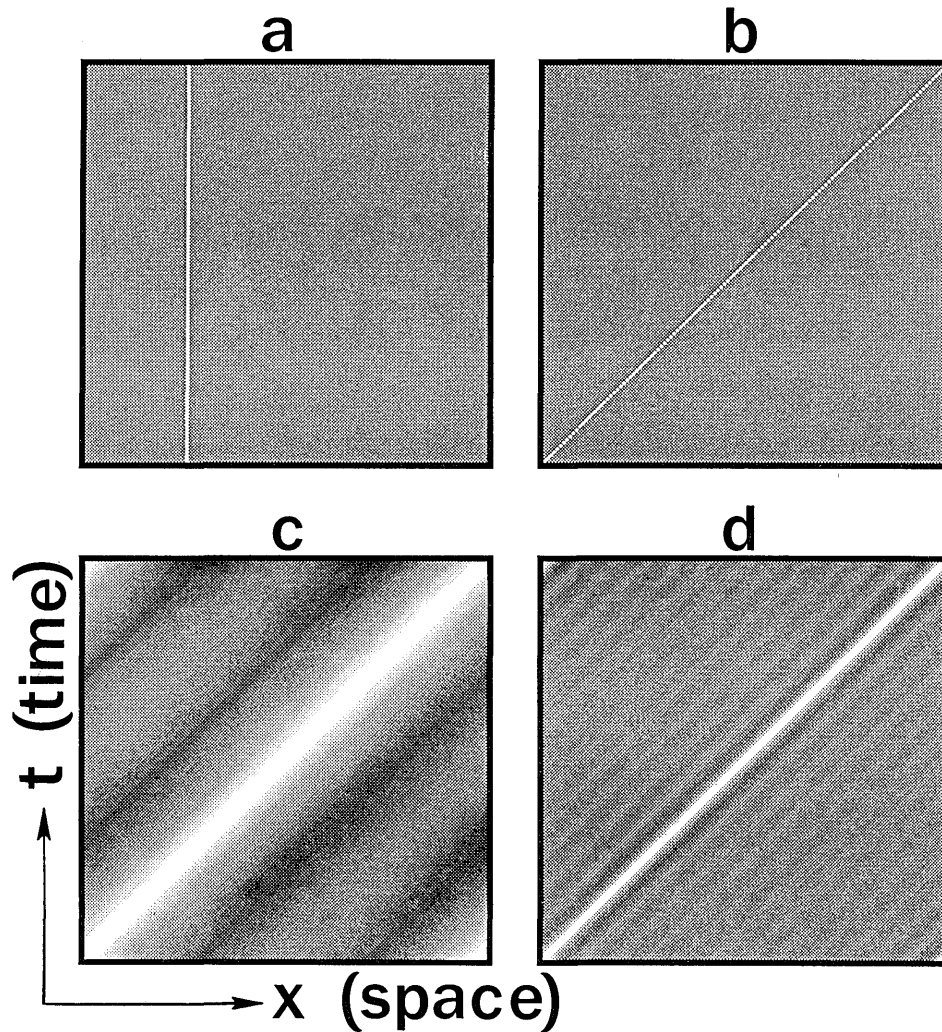


Fig. 1. Spatiotemporal Fourier analysis of a rightward-stepping bar. The abscissa represents horizontal space; the ordinate represents time. a, One frame of a movie of a rightward-stepping vertical bar. b, Horizontal-temporal cross section of a rightward-stepping vertical bar. c, Approximation to the rightward-stepping bar obtained by taking an equally weighted sum of $\{\cos(2\pi n(x/X - t/T))\}_{n=1, 2}$. d, Approximation to the rightward-stepping bar obtained by taking an equally weighted sum of $\{\cos(2\pi n(x/X - t/T))\}_{n=1, 2, \dots, 12}$.

or up and to the right). For example, Fig. 1a represents a single frame of a white vertical bar, extended up and down through the field of vision. Figure 1b shows the space-time representation of the bar in Fig. 1a, which appears at the left at time zero and moves at a constant rate to the right during the time course of the display.

For the moment, we shall generalize broadly, using the word sum to describe both finite and countable summations as well as integrations over bounded and unbounded real intervals. In this case, we can do approximate justice to some basic facts about visual stimuli and their Fourier transforms without getting bogged in technicalities. Any spatiotemporal stimulus l can be decomposed into a weighted sum of appropriately phase-shifted, drifting sinusoidal gratings. Moreover, this sum is unique: that is, there is only one assignment of weights and phases to drifting gratings that recaptures l in the corresponding sum.

Indeed, the Fourier transform of l is often defined to be the function that makes this assignment. There are, however, various other commonly encountered equivalent defini-

tions of the Fourier transform (one of which we shall shortly adopt) that may be more convenient for certain purposes.

Example: Fourier Components of a Rightward-Stepping Vertical White Bar

Most of the action of the moving bar stimulus l defined by Figs. 1a and 1b takes place along the line $L = \{(x, t) | x = t\}$ in Fig. 1b; that is, the points at which l deviates most from its mean value are along this line. For our purposes, the most useful indicator of where the action is in a given stimulus f is the squared deviation of f from its overall mean value at each point in its domain. As is clear, l deviates most energetically from its mean along the line L .

What spatiotemporal sinusoidal gratings are weighted most heavily in the Fourier sum yielding l ? A good way to answer this question is to ask another: What gratings can be shifted in phase so as to match l most closely? Those sinusoids that can be shifted so as to have high values where l has high values and low values where l has low values are the ones that will figure most heavily in the weighted sum com-

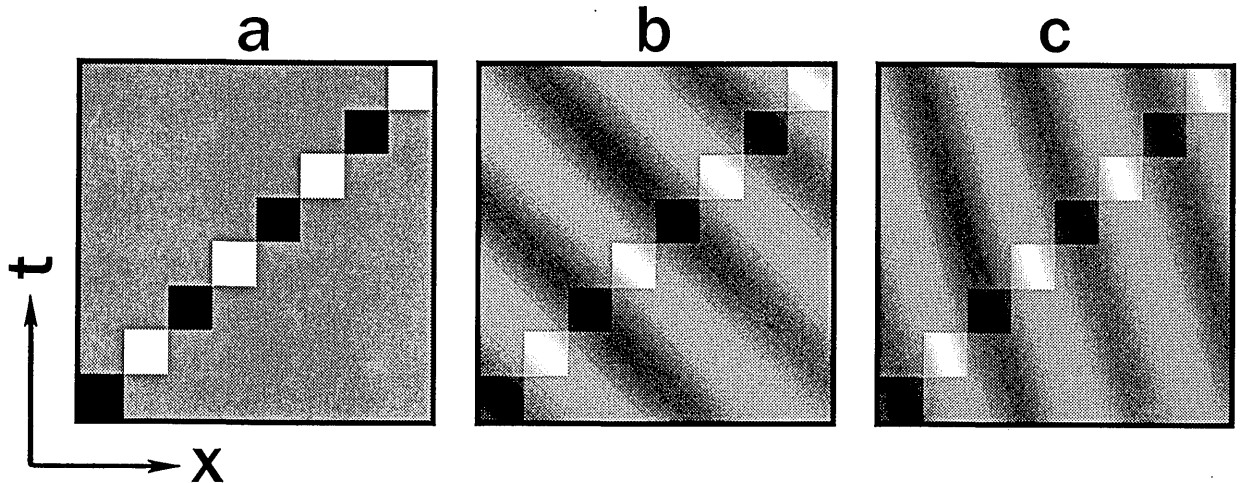


Fig. 2. Spatiotemporal Fourier analysis of stimulus h , a rightward-stepping, contrast-reversing vertical bar. a, Horizontal-temporal cross section of h . b, Horizontal-temporal cross section of a vertical, leftward-drifting sinusoid, which correlates well with h : $\cos(2\pi(2x/X + 2t/T) - \pi/2)$. c, Horizontal-temporal cross section of a more slowly leftward-drifting sinusoid, which also correlates well with h : $\cos(2\pi(3x/Xz + t/T) - \pi/2)$.

posing l . In short, those gratings that can be phase shifted so as to correlate best with l will have the highest amplitudes (weights) in the sum.

The sinusoidal gratings that correlate best with $l(x, t)$ of Fig. 1b are those that assume the value 1 along the line L , that is, all the sinusoids in the set

$$\Omega = \{\cos(\alpha x - \alpha t) | \alpha \in \mathbb{R}\}.$$

Figures 1c and 1d illustrate how l is approximated more and more closely by taking sums involving more and more (equally weighted) elements of Ω .

Example 1: Rightward-Stepping, Contrast-Reversing Vertical Bar

Contrast-reversing stimuli are critical for understanding the implications of Fourier analysis. Note first that, as in the case of l defined in Fig. 1, most of the power of h in Fig. 2 is centered along the line L . However, the elements of Ω contribute no power to h . To see this, note that the value of h flipflops around the mean luminance along L , while the value of any element $C \in \Omega$ remains constant; thus the value of the product of h with C will flipflop (with h) around the mean luminance over the points of L and will be zero everywhere else. Consequently, the sum taken over all points (x, t) of the product $h(x, t)C(x, t)$ is zero. This is equivalent to saying that the correlation of h with C is zero.

On the other hand, the function

$$C(x, t) = \cos(\alpha x + \beta t + \rho)$$

correlates positively with h when α and β are chosen so that the crests and troughs of C slope across L and oscillate at an appropriate frequency. ρ can then be chosen to lay the crests of C across the bright regions of h and the troughs across the dark regions. Examples of sinusoids that correlate well with h are given in Figs. 2b [$\cos(3x + t - \pi/2)$] and 2c [$\cos(2x + 2t - \pi/2)$].

Direction of Drift in Sinusoidal Gratings

For each nonnegative real number α , $\cos(\alpha x - \alpha t)$ drifts from left to right. By contrast, $\cos(\alpha x + \alpha t)$ drifts at the

same rate from right to left. For any $\omega, \tau, \rho \in \mathbb{R}$, if $\omega = 0$, the grating

$$C(x, t) = \cos(\omega x + \tau t + \rho)$$

has constant value over space but oscillates in time with frequency τ . Otherwise (if $\omega \neq 0$) C drifts with speed $|\tau/\omega|$; it drifts rightward if $\tau/\omega \leq 0$ and leftward if $\tau/\omega > 0$. Accordingly, we call C rightward drifting if $\tau/\omega < 0$, leftward drifting if $\tau/\omega > 0$, and stationary if $\tau = 0$.

3. THE MOTION-FROM-FOURIER-COMPONENTS PRINCIPLE

For any real-valued function, f , the sum (taken over all points in the domain of f) of the squared values of f is called the power in f . Parseval's relation states that the power in f is proportional to the sum of the squared amplitudes of the sinusoids into which f can be (uniquely) decomposed.

Thus, in particular, we can tally up the power in a dynamic visual stimulus either point by point in space-time or drifting sinusoid by drifting sinusoid. Of course, considering the unambiguous, uniform apparent motion displayed by drifting sinusoidal gratings, it would seem to make more sense for a motion-perception system to do its power accounting across the sinusoids composing the stimulus.

These considerations lead naturally to a commonly encountered general rule for predicting the apparent motion of an arbitrary horizontal stimulus $l(x, t)$: For l considered as a linear combination of sinusoidal gratings, compare the power in l of the rightward-drifting gratings with the power of the leftward-drifting gratings; if most of l 's power is contributed by rightward-drifting gratings, then perceived motion should be to the right. If most of the power resides in the leftward-drifting gratings, perceived motion should be to the left. Otherwise l should manifest no decisive motion in either direction.

This prediction rule for horizontally moving stimuli is a restricted version of the motion-from-Fourier-components (MFFC) principle: More generally, let L be any visual stimulus; that is, $L: X \times Y \times T \rightarrow \mathbb{R}$, for bounded real intervals X, Y , and T , where for any $(x, y, t) \in X \times Y \times T$, $L(x, y, t)$ is

construed as the luminance of a point (x, y) in a visual field at time t . A more general version of the MFFC principle is as follows: For L to exhibit motion in a certain direction in the neighborhood of some point $(x, y, t) \in \mathbb{R}^3$, there must be some spatiotemporal volume Δ in some sense proximal to (x, y, t) such that the Fourier transform of L computed locally across Δ has substantial power over some regions of the frequency domain whose points correspond, in the space-time domain, to sinusoidal gratings whose direction of drift is consonant with the motion perceived.

That any standard version of the MFFC principle cannot account for all phenomena associated with human motion perception was demonstrated by Sperling,¹³ who described the following three-flash stimulus. Frame 0 is a rectangular block of contiguous small squares, each of which is independently painted black or white with equal probability. In frame 1, a subblock B_1 of frame 0 is scrambled (that is, in frame 1, each component square within B_1 is independently repainted black or white with equal probability). In frame 2 a different subblock, B_2 , is scrambled. For many sizes of rectangles and frame presentation rates, such a stimulus elicits apparent motion in the direction from B_1 to B_2 ; nonetheless, it is unlikely to correlate significantly with any given spatiotemporal sinusoidal grating.

It is our purpose here to build on these observations. We shall first give precise formulation to the notion of a random stimulus and then define a certain class of random stimuli (the class of drift-balanced random stimuli) that is useful in studying visual perception (since any motion displayed by a drift-balanced random stimulus cannot be explained in terms of the MFFC principle). We proceed to show that the (spatiotemporal) convolution of two drift-balanced random stimuli is drift balanced and mention some of the psychophysical implications of this fact. In proposition 3 below we prove that linear combinations of certain drift-balanced random stimuli are themselves drift balanced (this result, which is illustrated with a variety of stimulus examples, is particularly useful in constructing drift-balanced random stimuli that display consistent apparent motion across independent realizations). In Section 7 we provide an alternative characterization of the class of drift-balanced random stimuli in terms of simple point-delay Reichardt detectors (or autocorrelation coefficients) and apply this characterization to distinguish the subclass of drift-balanced random stimuli that we call microbalanced. A random stimulus I is microbalanced if, for any space-time-separable function W , the result WI of windowing I by W is drift balanced. We derive a collection of basic results about microbalanced random stimuli and show that, in fact, all the demonstration stimuli previously defined (demonstrations 1–5 below) are microbalanced. Among other things, we prove that the expected response of any elaborated Reichardt detector^{1,2} to any microbalanced random stimulus is zero at any instant in time. Finally, we observe some salient psychophysical properties of microbalanced random stimuli and discuss some of the possible explanations of the non-Fourier motion elicited by such stimuli.

4. PRELIMINARIES

In this paper we deal with properties of random stimuli. Roughly speaking, a random stimulus is a jointly distributed family of random variables assigned to a grid of locations

covering the visual field across time. In this section we collect the tools appropriate for dealing with such objects. This section is split into two subsections, one devoted to continuous random variables, in which we introduce explicitly some notation for handling integration and define a density; and one devoted to discrete dynamic visual stimuli and their Fourier transforms, in which we identify a stimulus [an assignment of luminance (nonnegative, real values) to a regular grid of points throughout visual space and time] with its contrast modulation function (the normalized deviation of luminance from its mean) and introduce frequency-domain notation.

Continuous Random Variables

Our stimuli are real-valued, randomly varying functions of a discrete domain. The luminances assigned to points (pixels) are, in general, jointly distributed random variables. The basic definitions and proofs that we present here presuppose that these random variables are real valued and continuous. (In general, the discrete-case analogs are simpler and should be obvious.)

Let \mathbf{Z} (\mathbf{Z}^+) denote the set of integers (positive integers), and let \mathbb{R} (\mathbb{R}^+) denote the real (positive real) numbers.

The following conventions are useful. As usual, call any subset $\alpha \subseteq \mathbb{R}$ an interval if and only if (iff), for any $x, z \in \alpha$ and any $y \in \mathbb{R}$, if $x \leq y \leq z$, then $y \in \alpha$; more generally, for any $k \in \mathbf{Z}^+$, call any subset $\beta \subseteq \mathbb{R}^k$ an interval of \mathbb{R}^k iff β is the Cartesian product of (possibly unbounded) real intervals $\beta_0, \beta_1, \dots, \beta_{k-1}$. In this case, for any function $f: \mathbb{R}^k \rightarrow \mathbb{R}$, it is convenient to indicate the integral of f over β , if it exists, as

$$\int_{\beta} f(\nu) d\nu.$$

Moreover, we call any nonnegative, real-valued function f of \mathbb{R}^k a density iff f is integrable over \mathbb{R}^k and

$$\int_{\mathbb{R}^k} f(\nu) d\nu = 1.$$

Discrete Dynamic Visual Stimuli and Their Fourier Transforms

Contrast Modulation

Luminance is physically constrained to be a nonnegative quantity. Psychophysically, however, the significant quantity is contrast, the normalized deviation at each time t of luminance at each point (x, y) in the visual field from a base level, or level of adaptation, which reflects the average luminance over points proximal to (x, y, t) in space and time. We shall restrict our attention throughout this paper to stimuli for which it may be assumed that the base luminance level μ is uniform over the significant spatiotemporal locations in the display. In practice, this condition is met if (i) subjects are adapted sufficiently to a field of uniform luminance μ before the onset of non- μ luminances and (ii) the duration over which non- μ luminances are displayed is sufficiently brief.

For any stimulus L with base luminance μ , call the function I satisfying

$$L = \mu(1 + I) \quad (1)$$

the *contrast modulator* of L (and note that $I \geq -1$).

Psychophysically, it is well known that, over substantial ranges of μ , the apparent motion of L does not depend on μ .

Thus the contrast modulator I of L emerges as a likely function to analyze for the motion information carried by L . Accordingly, we shall shift our focus from luminance to contrast and identify L with its contrast modulator, dropping reference to adaptation level.

Specifically, we shall call any function $I:Z^3 \rightarrow \mathbb{R}$ a *stimulus* iff $I[x, y, t] = 0$ for all but finitely many points $(x, y, t) \in Z^3$.

Strictly speaking, we should also require that I never drop below -1 . This restriction, however, would lead to unnecessary complications in dealing with various sorts of combinations of stimuli. In all cases, the points that we wish to make tolerate rescaling of stimuli by arbitrary multiplicative constants to settle their minimal values to some perceptually appropriate level between -1 and 0 . Accordingly, we drop the restriction that $I \geq -1$.

In general, we shall consider stimuli of two spatial dimensions and time. The reader may find it convenient to think of the first spatial dimension (which we shall always index by x) as horizontal, with indices increasing to the right, and the second spatial dimension (always indexed by y) as vertical, with indices increasing upward. The temporal dimension is always indexed by t .

Frames and Frame Blocks

For any stimulus I , we call the restriction of I to $Z^2 \times \{t\}$ the t th *frame* of I . In all the stimulus examples that we shall consider, frames clump into blocks: specifically, for each demonstration stimulus I defined in this paper, there are integers k and N such that all changes in luminance occur in frames kn , where $n = 0, 1, \dots, N$, and otherwise luminance remains constant across frames. The group of identical frames between and including frames kn and $kn + k - 1$ we shall call the n th *frame block* of I .

Any stimulus I is nonzero at only a finite number of points in its countably infinite domain. Consequently, (i) the mean value of I is 0 , and (ii) the power in I is finite.

From property (ii) we observe that I has a well-defined Fourier transform, which we denote by \bar{I} . Specifically,

$$\bar{I}(\omega, \theta, \tau) = \sum_{(x,y,t) \in Z^3} I[x, y, t] \exp(-j(\omega x + \theta y + \tau t))$$

(analysis).

We shall always use square brackets around the arguments of discrete functions and parentheses around the arguments of continuous functions. Although \bar{I} is defined for all $(\omega, \theta, \tau) \in \mathbb{R}^3$, it is periodic over 2π in each variable. This fact is reflected in the inverse transform:

$$I[x, y, t] = \frac{1}{(2\pi)^3} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \bar{I}(\omega, \theta, \tau) \times \exp(j(\omega x + \theta y + \tau t)) d\omega d\theta d\tau$$

(synthesis).

In the Fourier domain we shall consistently use ω to index frequencies relative to x , θ to index frequencies relative to y , and τ to index frequencies relative to t . This convention is exemplified by the definition of \bar{I} above.

We distinguish the stimulus 0 by setting $0[x, y, t] = 0$ for all $x, y, t \in Z$. In parallel, we let $\bar{0}$ assign 0 to all $(\omega, \theta, \tau) \in \mathbb{R}^3$.

5. DRIFT-BALANCED RANDOM STIMULI

We begin by generalizing the notion of a stimulus to that of a random stimulus. Whereas a nonrandom stimulus assigns fixed values to Z^3 , a random stimulus I assigns jointly distributed random variables that deviate from zero at only a finite number of points.

Various expectations associated with I are defined easily. We shall be particularly interested in the expected power of \bar{I} at some point, (ω, θ, τ) in the frequency domain: $E[|\bar{I}(\omega, \theta, \tau)|^2]$. This reflects the expected power in I of a sinusoid C that modulates contrast at the rate of $\omega/2\pi$ cycles per column, $\theta/2\pi$ cycles per row, and $\tau/2\pi$ cycles per frame. The sinusoid with the same spatial frequency as C and moving at the same rate but in the opposite direction is obtained simply by reversing the direction of C 's temporal contrast modulation: that is, by modulating contrast $-\tau/2\pi$ cycles per frame. When the expected power in I of any given drifting sinusoid is matched by the expected power of the sinusoid of the same spatial frequency drifting at the same rate in the opposite direction, we call I drift balanced.

Although the MFFC principle suggests that drift-balanced random stimuli should not display consistent apparent motion across independent realizations, we shall provide examples of drift-balanced random stimuli (in Section 6) that do in fact display strong, consistent motion across trials.

Beyond these basic developments, two propositions are proved in this section. In proposition 1 we demonstrate that any random stimulus separable in space and time (see definition 3 below) is drift balanced, and in proposition 2 we show that the (spatiotemporal) convolution of any two independent, drift-balanced random stimuli is drift balanced.

We now proceed more precisely as follows.

Definition 1: Random Stimulus

Call any family $I[x, y, t], (x, y, t) \in Z^3$, of random variables jointly distributed with density f , a *random stimulus* when

- (i) $I[x, y, t] = 0$ for all but a finite subset $\alpha \subset Z^3$ and
- (ii) $E[I[x, y, t]^2]$ exists for all $(x, y, t) \in Z^3$.

Expectations Related to I

With k the cardinality of α , we set up a one-to-one correspondence between dimensions of \mathbb{R}^k and points of α so that each coordinate of any vector $i \in \mathbb{R}^k$ corresponds to one of the points of α . We can now treat i as a stimulus (whose nonzero values are restricted to the points of α). In particular, letting $i_{(p,q,r)}$ denote the coordinate of i corresponding to a given $(p, q, r) \in \alpha$, we set

$$i[x, y, t] = \begin{cases} i_{(x,y,t)} & \text{if } (x, y, t) \in \alpha \\ 0 & \text{otherwise} \end{cases}$$

for any $(x, y, t) \in Z^3$. We can now conveniently formulate various expectations associated with I ; in particular, we define the expectation of I by

$$E_I[x, y, t] = \int_{\mathbb{R}^k} i[x, y, t] f(i) di$$

for all $(x, y, t) \in Z^3$. (Note that E_I is a nonrandom stimulus.) Consider the Fourier transform of E_I :

$$\begin{aligned} \bar{E}_I(\omega, \theta, \tau) &= \sum_{(x,y,t) \in \mathbb{Z}^3} \int_{\mathbb{R}^k} i[x, y, t] f(i) di \\ &\quad \times \exp(-j(\omega x + \theta y + \tau t)) \\ &= \int_{\mathbb{R}^k} \sum_{(x,y,t) \in \mathbb{Z}^3} i[x, y, t] \exp(-j(\omega x + \theta y + \tau t)) f(i) di \\ &= \int_{\mathbb{R}^k} \bar{i}(\omega, \theta, \tau) f(i) di = E[\bar{I}(\omega, \theta, \tau)]. \end{aligned}$$

This leads to the following observation.

Observation 1

The Fourier transform of the expectation of a random stimulus I is equal to the expectation of the Fourier transform of I .

Note especially, here, the implication that $E_I = 0$ iff $E[\bar{I}] = \bar{0}$.

We call any random stimulus I invariant iff there exists a stimulus S such that $I = S$ with probability 1.

Example 2: Randomly Contrast-Reversing, Rightward-Stepping Vertical Bar

Let the random stimulus I contain four frame blocks indexed 0, 1, 2, and 3, and let each frame block be composed of a horizontal sequence of four rectangles indexed 0, 1, 2, and 3 from left to right. Let $\phi_0, \phi_1, \phi_2,$ and ϕ_3 be pairwise independent random variables, each taking the value C or $-C$ with equal probability. Give rectangle i in frame block i the value assumed by ϕ_i , and give all other pixels the value 0.

The restriction of I to any one of its rows is characterized by Fig. 3, as a function of x along the horizontal axis and t along the vertical axis. As is clear, for any $(x, y, t) \in \mathbb{Z}^3$,

$$E[I[x, y, t]] = 0;$$

that is, $E_I = 0$, from which we infer that $\bar{E}_I = \bar{0}$.

An interesting fact that may not be so obvious, however, (this follows from corollary 1 below) is that the expected power contributed to I by any given drifting sinusoidal grating is equal to the expected power contributed by the grating of the same spatial frequency drifting at the same rate in the opposite direction. This may seem surprising in light of the MFFC principle, since any realization of I is marked by a systematic, left-to-right perturbation across time, which (as one might expect) tends, under appropriate viewing conditions, to be perceived as motion from left to right. Indeed, as we shall see in Section 6, it is quite easy to construct random stimuli with this property that nonetheless display striking, reliable apparent motion in a fixed direction.

This fact motivates a notion central to this paper: that of a drift-balanced random stimulus (see definition 2 below). As the name suggests, a drift-balanced random stimulus is one for which the expected contribution of any given drifting sinusoidal grating is balanced by (equal to) the expected contribution of the corresponding grating drifting at the same rate in the opposite direction. Of course, just as a given random variable may have little or no probability of assuming a value equal to its expectation, a particular real-

frame block 3	0	0	0	ϕ_3
frame block 2	0	0	ϕ_2	0
frame block 1	0	ϕ_1	0	0
frame block 0	ϕ_0	0	0	0
rectangle	0	1	2	3

Fig. 3. Rightward-stepping, randomly contrast-reversing vertical bar: a horizontal-temporal diagram of the random stimulus I , a vertical bar that appears with contrast C or $-C$ randomly assigned and steps its width rightward three times over a zero-contrast visual field, assuming contrast C or $-C$ with equal probability with each step. The expected power in I of any given drifting sinusoid is equal to the expected power of the sinusoid of the same spatial frequency drifting at the same rate but in the opposite direction.

ization of a drift-balanced random stimulus, I , does not, in general, have perfectly balanced components. However, when gauged over a number of independent realizations, the mean contribution of a particular Fourier component of I tends to balance against the contribution of the corresponding, oppositely moving component.

Definition 2: Drift-Balanced Random Stimulus

Call any random stimulus I drift balanced iff, for any $\omega, \theta, \tau \in \mathbb{R}$,

$$E[|\bar{I}(\omega, \theta, \tau)|^2] = E[|\bar{I}(\omega, \theta, -\tau)|^2]. \tag{2}$$

[For a proof that the expectations in Eq. (2) exist, see Appendix A.] Notice that, because I is real valued, Eq. (2) is equivalent to

$$E[|\bar{I}(\omega, \theta, \tau)|^2] = E[|\bar{I}(-\omega, -\theta, \tau)|^2];$$

that is, I is drift balanced iff the expected power in I of any given drifting sinusoidal grating is equal to the expected power of the grating with the same spatial frequency drifting at the same rate but in the opposite direction.

As we shall see in Section 6, the following class of random stimuli is useful in constructing drift-balanced random stimuli that display consistent motion.

Definition 3: Space-Time-Separable Random Stimulus

Call any random stimulus I space-time separable iff, for any $(x, y, t) \in \mathbb{Z}^3$,

$$I[x, y, t] = g[x, y]h[t],$$

for jointly distributed real random functions g and h .

Immediately we note a simple proposition.

Proposition 1

Any space-time-separable random stimulus is drift balanced.

Proof

Let I be a space-time-separable random stimulus, with

$$I[x, y, t] = g[x, y]h[t]$$

for all $(x, y, t) \in \mathbf{Z}^3$; then

$$|\bar{I}(\omega, \theta, \tau)|^2 = |\bar{g}(\omega, \theta)|^2 |\bar{h}(\tau)|^2.$$

Thus, since h is real valued,

$$|\bar{I}(\omega, \theta, \tau)|^2 = |\bar{g}(\omega, \theta)|^2 |\bar{h}(-\tau)|^2 = |\bar{I}(\omega, \theta, -\tau)|^2.$$

Taking expectations of both sides yields Eq. (2). ■

It would be surprising for any space-time-separable random stimulus I to exhibit strong, consistent motion in a fixed direction, since the only sort of temporal contrast change induced by I is a spatially global modulation.

However, as we have hinted in example 1, there do exist drift-balanced random stimuli that exhibit decisive motion in a fixed direction not only on the average across a number of trials but on virtually each display. In Section 6 we shall provide some general results that are useful for constructing a broad range of drift-balanced random stimuli that show strong motion. However, we shall show first that the spatio-temporal convolution of independent drift-balanced random stimuli is drift balanced and briefly mention some of the ramifications of this fact.

Proposition 2

The (spatiotemporal) convolution of independent, drift-balanced random stimuli is drift balanced.

Proof

Let I and J be independent drift-balanced random stimuli. For any random stimuli we have

$$|\overline{I * J}|^2 = |\bar{I}|^2 |\bar{J}|^2.$$

The independence of I and J implies that

$$E[|\overline{I * J}|^2] = E[|\bar{I}|^2] E[|\bar{J}|^2].$$

Thus, since I and J are drift balanced, we find that, for any $\omega, \theta, \tau \in \mathbf{R}$,

$$\begin{aligned} E[|\overline{I * J}(\omega, \theta, \tau)|^2] &= E[|\bar{I}(\omega, \theta, \tau)|^2] E[|\bar{J}(\omega, \theta, \tau)|^2] \\ &= E[|\bar{I}(\omega, \theta, -\tau)|^2] E[|\bar{J}(\omega, \theta, -\tau)|^2] \\ &= E[|\overline{I * J}(\omega, \theta, -\tau)|^2]. \end{aligned}$$

Most computational models of the sensory transformations mediating human perception routinely apply a spatio-temporal, linear, shift-invariant filter to the input stimulus. The impulse response (i.e., convolution kernel) of any such filter can, of course, be regarded as an invariant stimulus. Typically the filters applied are drift balanced.^{1-4,17} Obviously, filters that depend on only spatial characteristics of the stimulus being processed are drift balanced (for instance, all manner of oriented, band-tuned, spatial edge detectors). Similarly, filters (such as flicker detectors) that depend on only temporal stimulus characteristics are drift balanced. More generally, all space-time-separable filters are drift balanced (proposition 1). Thus, given a drift-balanced random input stimulus, the output of many of the filters that are commonly thought to function in the early stages of human visual processing is also drift balanced.

6. CONSISTENT APPARENT MOTION FROM DRIFT-BALANCED STIMULI

We begin this section by noting some general results concerning linear combinations of random stimuli, leading up to proposition 3 below, in which we show that any linear combination of pairwise independent, drift-balanced random stimuli, all of which have expectation 0, is drift balanced. (Actually, this is an implication of proposition 3, which is slightly more general.) From this finding follow corollaries 1 and C1 (C1 in Appendix C), each of which gives rise to specific examples of drift-balanced random stimuli that elicit consistent apparent motion. Several of these examples are detailed in this section. Experimental findings with regard to these example random stimuli are reported.

One may wonder whether linear combinations of independent drift-balanced random stimuli are drift balanced. That this is not the case is evident from the fact that any invariant stimulus whatsoever can be expressed as a linear combination of shifted impulses, which are, of course, jointly independent and individually drift balanced.

Although linear combinations of arbitrary, pairwise independent, drift-balanced random stimuli are not generally drift balanced, if we impose an additional constraint on the random stimuli to be summed we can ensure that the resultant linear combination is indeed drift balanced.

The following lemma bears on this issue.

Lemma 1

Let S be a random stimulus equal to the sum of a set Ω of pairwise independent random stimuli; then

$$E[|S|^2] = |\bar{E}_s|^2 + \sum_{I \in \Omega} E[|N_I|^2],$$

where $N_I = I - E_I$ for each $I \in \Omega$.

Proof

See Appendix B.

Immediately we note a useful result concerning linear combinations of drift-balanced random stimuli:

Proposition 3

Let $\Omega = \Theta \cup \{I\}$ be a set of pairwise independent, drift-balanced random stimuli, such that I is invariant and each member of Θ has an expectation of 0. Then any linear combination, S , of the elements of Ω is drift balanced.

Proof

A drift-balanced random stimulus rescaled by a constant is drift balanced. Thus we assume with no loss of generality that S is just a sum of pairwise independent drift-balanced random stimuli.

Note that (i) $I = E_I$ (hence $N_I = I - E_I = 0$) and (ii) for all $J \in \Theta, N_J = J - E_J = J$. Thus from lemma 1 we observe for any $\omega, \theta, \tau \in \mathbf{R}$

$$\begin{aligned}
E[|\bar{S}(\omega, \theta, \tau)|^2] &= |\bar{E}_S(\omega, \theta, \tau)|^2 + \sum_{J \in \Omega} E[|\bar{N}_J(\omega, \theta, \tau)|^2] \\
&= |\bar{I}(\omega, \theta, \tau)|^2 + \sum_{J \in \Theta} E[|\bar{J}(\omega, \theta, \tau)|^2] \\
&= |\bar{I}(\omega, \theta, -\tau)|^2 + \sum_{J \in \Theta} E[|\bar{J}(\omega, \theta, -\tau)|^2] \\
&= E[|\bar{S}(\omega, \theta, -\tau)|^2]. \quad \blacksquare
\end{aligned}$$

Note, in particular, that this result holds for $I = 0$.

As is reasonably clear from proposition 3 (since space-time-separable random stimuli are drift balanced), any sum of pairwise independent, space-time-separable random stimuli, all with an expectation of $\mathbf{0}$, is drift balanced. In corollary 1 this principle is applied to generate a class of drift-balanced random stimuli, certain instances of which exhibit strong, consistent apparent motion in a fixed direction.

Corollary 1

For $M \in \mathbf{Z}^+$, let $\phi_0, \phi_1, \dots, \phi_{M-1}$ be pairwise independent random variables, each with expectation 0; and, for $m = 0, 1, \dots, M-1$, let $f_m: \mathbf{Z}^2 \rightarrow \mathbb{R}$ and $g_m: \mathbf{Z} \rightarrow \mathbb{R}$, and let the product $f_m g_m$ be 0 at all but finitely many points of \mathbf{Z}^3 ; then the random stimulus I defined by setting

$$I[x, y, t] = \sum_{m=0}^{M-1} \phi_m f_m[x, y] g_m[t], \quad (3)$$

is drift balanced.

The proof is obvious from propositions 1 and 3.

A simple yet compelling counterexample to the MFFC principle may now be constructed as follows.

Demonstration 1: A Randomly Contrast-Reversing, Rightward-Stepping Rectangle

For some $M \in \mathbf{Z}^+$, let the random stimulus I be composed of M frame blocks indexed $0, 1, \dots, M-1$, and let each frame block be composed of a horizontal sequence of M rectangles indexed $0, 1, \dots, M-1$ from left to right (see example 2 and Fig. 3). Let $\phi_0, \phi_1, \dots, \phi_{M-1}$ be pairwise independent random variables, each taking the value C or $-C$ with equal probability. Give rectangle i in frame block i the value assumed by ϕ_i , and give all other pixels the value 0. We can now define I by Eq. (3) by letting $f_m[x, y]$ take the value 1 in the m th rectangle and 0 everywhere else and letting $g_m[t]$ take the value 1 in the m th frame block and 0 everywhere else.

The apparent motion of this stimulus is quite easy to imagine: throughout frame block 0, rectangle 0 is present on the left-hand side of the stimulus field; it is assigned contrast of C or $-C$ with equal probability. In frame block 1, rectangle 0 turns off (goes to contrast 0), and rectangle 1, abutting rectangle 0 on the right, turns on, again with contrast C or $-C$ assigned with equal probability, independent of the contrast of the first rectangle. In each successive frame block, one rectangle turns off, and a new rectangle turns on directly to the right of its predecessor, with contrast either C or $-C$, independent of any other rectangle.

Figure 4a displays a realization of one version of the random stimulus I defined in demonstration 1 with $M = 8$. This random stimulus and others that we shall discuss were tested experimentally on two subjects. Before discussing responses to I in particular, we describe the experimental arrangements for these observations.

General Method

We describe here the procedure for demonstrations 1 (stimulus I), 2 (K), 3 (J), 4 (H), and 5 (G). All stimulus presentations were made on a Conrac 7211 RGB monitor driven by an Adage graphics display processor. The display area was $28 \text{ cm} \times 32 \text{ cm}$, and displayed intensities were greenish white. The spatial resolution was 512×512 pixels, the temporal resolution was 60 frames/sec, and the intensity resolution was 256 gray levels.

Two subjects were involved in each of the studies: CC (the experimenter) and DY (a naive subject). For each demonstration, each subject viewed 30 independent realizations of the random stimulus. On each presentation, the non-Fourier motion of the stimulus (I, K, J, H , or G) was left to right or right to left with equal probability. For instance, I 's randomly contrast-reversing rectangle stepped left to right or right to left with equal probability.

Subjects adapted before each session to a uniform screen of luminance 80 cd/m^2 ; other luminances were linearized carefully relative to the mean. All stimuli were viewed foveally and binocularly, from a distance of 2 m. On each trial, a central cue spot ($0.5 \text{ deg} \times 0.5 \text{ deg}$) of low positive contrast came on 2 sec before the onset of the stimulus and disappeared 1 sec before the onset. Subjects were instructed to maintain their gazes throughout the trial on the cue spot point and were required to indicate the predominant direction of apparent motion (left or right) by entering either an L or an R on a terminal keyboard.

Method for Demonstration 1

In the version of I viewed by our subjects, frame blocks lasted $1/60$ sec; spatial rectangles measured approximately 2 deg (horizontal) $\times 2 \text{ deg}$ (vertical) and $C = 0.25$. The contrast of 0.25 was chosen because it produced easily visible motion and yet was small enough that psychophysical, as well as physical, equivalence of positive and negative increments was likely to hold.

Results

Subject CC (DY) reported apparent motion in the step direction on 30 (29) of 30 trials.

Discussion

The essential trick of the rightward stepping bar was to modulate the contrast (that is, the absolute deviation from zero) of a field of static, spatially independent, zero-mean noise as a function of space and time. This notion of spatio-temporal modulation of contrast needs some explanation. Let J be a random stimulus with expectation $\mathbf{0}$, let W be a nonnegative function of \mathbf{Z}^3 (space and time), and consider $I = WJ$. In general, J 's distance from $\mathbf{0}$, be it positive or negative, is magnified (or damped) by W 's value at each point in space and time. Thus I is obtained by letting W modulate the (absolute) contrast of J .

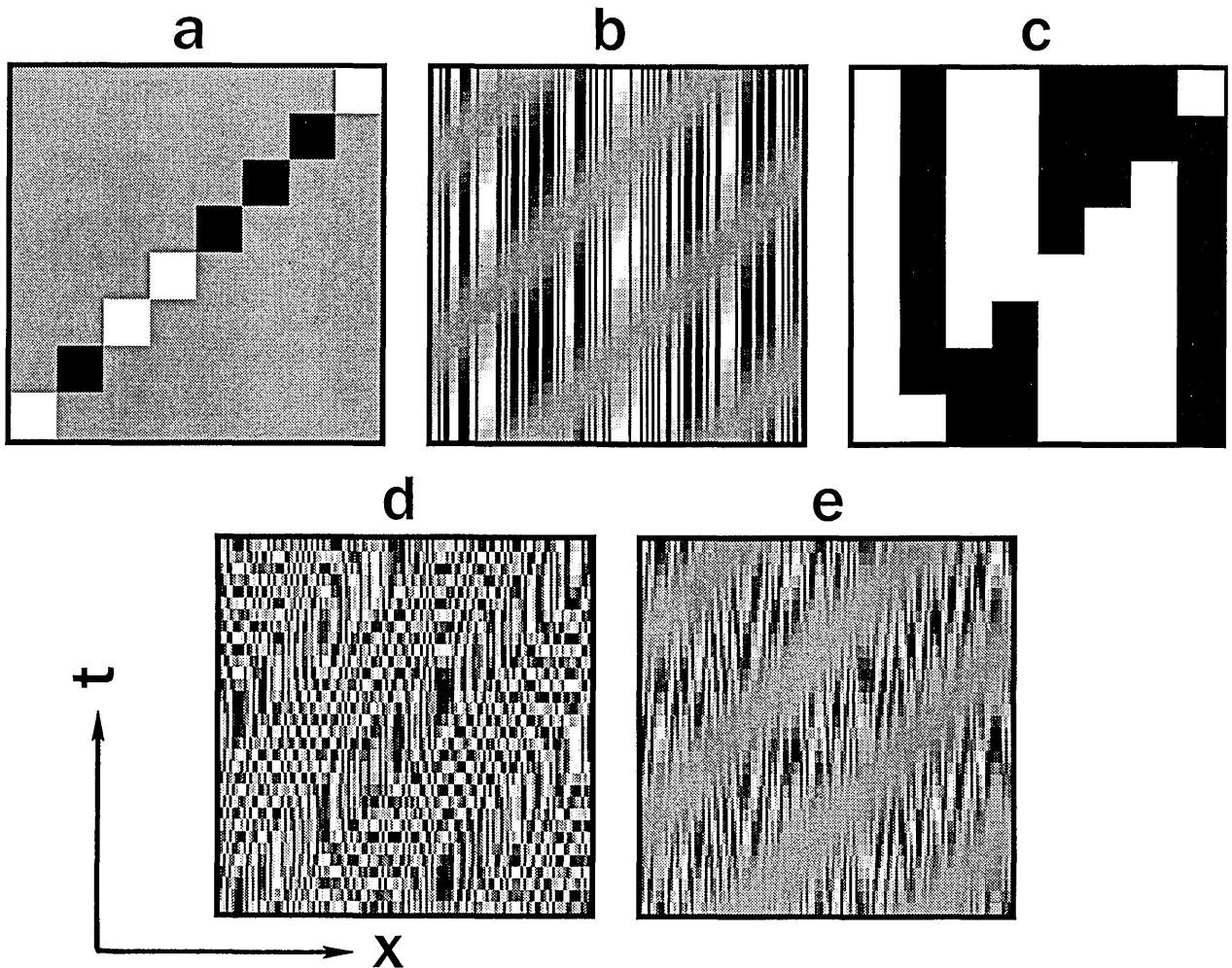


Fig. 4. a, Rightward-stepping, randomly contrast-reversing vertical bar: a horizontal-temporal cross section of a realization of the random stimulus I (see demonstration 1). I is the sum of pairwise independent space-time-separable random stimuli, each of which has an expectation of 0; consequently I is drift balanced (by corollary 1). b, Modulation of the contrast of a static noise field by a drifting sinusoidal grating: a horizontal-temporal cross section of a realization of the random stimulus K (demonstration 2). That K is drift balanced follows from corollary 1. c, Traveling contrast reversal of a noise field: a horizontal-temporal cross section of a realization of the random stimulus J (demonstration 3). J is the sum of pairwise independent space-time-separable random stimuli, each of which has an expectation of 0 and is thus drift balanced (by corollary 1). Note that, in contrast to $|I|$ (for I of Fig. 4a), $|J|$ is devoid of motion information. d, Modulation of the flicker frequency of a flickering noise field by a drifting grating: a horizontal-temporal cross section of a realization of the random stimulus H (demonstration 4). That H is drift balanced is a consequence of corollary C1 (in Appendix C). The motion of H is derived from spatiotemporal modulation of the frequency of sinusoidal flicker, where the phase of the flicker is random over space. e, Modulation of the contrast of a flickering noise field by a drifting sinusoidal grating: a horizontal-temporal cross section of a realization of the random stimulus G (demonstration 5). G is drift balanced (by corollary C1). The motion of G is derived from spatiotemporal modulation of the amplitude of sinusoidal flicker, where the flicker phase is random over space.

To see how this notion applies to I of demonstration 1, note that we can look at I as the result of multiplying a field J of random black or white rectangles persisting through M chunks of time by a function W , which (for $m = 0, 1, \dots, M - 1$) is 1 in the m th frame block for the points in the m th rectangle from the left and 0 everywhere else.

Elaborations of this basic contrast-modulation scheme are easy to construct. Consider, for instance, demonstration 2.

Demonstration 2: Contrast Modulation of a Static Noise Field by a Drifting Sinusoid

We compose the random stimulus K of N frame blocks, each containing a horizontal row of rectangles, indexed $0, 1, \dots, M - 1$ from left to right. For $m = 0, \dots, M - 1$, let $f_m[x, y]$ take the value 1 in the m th rectangle and 0 elsewhere, and let

$g_m[t]$ vary as a sinusoidal function of m and the frame block. Specifically, for each frame t in the n th frame block, let

$$g_m[t] = \frac{\cos[2\pi(\alpha m/M - \beta n/N)] + 1}{2}$$

for some spatial and temporal frequencies α and β . Let $\phi_0, \phi_1, \dots, \phi_{M-1}$ be pairwise independent random variables taking the values C and $-C$ with equal probability, for some contrast C , and define K by Eq. (3).

Whereas I of demonstration 1 merely picks out successive rectangles of spatial noise (independently assigned contrast C or $-C$) in successive time intervals, K is marked by high-power crests (α per frame block) separated by zero-power (gray) troughs sweeping at a constant rate from left to right over the row of rectangles, each of random contrast C or $-C$.

Figure 4b shows a realization of K , with $M = 128$, $N = 32$, and $\alpha = \beta = 2$.

Method

In the version of K viewed by our subjects, frame blocks lasted 1/60 sec, rectangles measured approximately (1/8 deg horizontal) by (2 deg vertical), and contrast $C = 0.25$.

Results

The cosine grating modulating the contrast of K was rightward or leftward drifting with equal probability. Subject CC (DY) reported apparent motion in the direction of drift in 30 (26) of 30 trials.

It might be that humans extract the motion from stimuli such as I (Fig. 4a) and K (Fig. 4b) simply by performing a Fourier power analysis on a rectified version of the stimulus. For instance, if subjects were able either (i) to disregard (set to 0) all negative contrast values or (ii) to map all contrasts onto their absolute values, then it is clear that a Fourier power analysis of the resultant rectified signal would correspond quite well to perceived motion. This explanation does not account for responses to stimuli of the type considered in demonstration 3.

Demonstration 3: Traveling Contrast Reversal of a Random Bar Pattern

Let $M \in \mathbf{Z}^+$. We construct the random stimulus J of $M + 1$ frame blocks indexed $0, 1, \dots, M$, each of which contains M rectangles indexed $0, 1, \dots, M - 1$ from left to right. Let $f_m[x, y]$ take the value 1 in the m th rectangle and zero elsewhere; let $g_m[t]$ be 1 in frame blocks 0 through m , -1 in frame blocks $m + 1$ through M , and 0 everywhere else. Let the random variables $\phi_0, \phi_1, \dots, \phi_{M-1}$ be pairwise independent, each taking a contrast value of C or $-C$ with equal probability, and use Eq. (3) to define J .

In frame block 0 of J , all M rectangles turn on, some with contrast C and others with contrast $-C$. In successive frame blocks $m = 1, 2, \dots, M$, exactly one of the rectangles changes contrast: the $(m - 1)$ th switches to C if its previous contrast was $-C$; otherwise it flips from C to $-C$. In frame block 1, the leftmost (0th) rectangle flips contrast; in frame block 2, rectangle 1 flips, and in successive frame blocks, successive rectangles flip contrast from left to right, until the $(M - 1)$ th rectangle flips in frame block M , after which all the rectangles turn off.

Method

The version of J viewed by subjects CC and DY contained nine frame blocks, each of which lasted 1/60 sec and contained eight spatial rectangles, each measuring approximately 2 deg \times 2 deg; $C = 0.25$.

Results

CC (DY) reported apparent motion in the direction traveled by the contrast flip in 30 (25) of 30 trials.

The next two stimuli (G of demonstration 4 and H of demonstration 5) are both drift balanced. The proof of this fact depends on a corollary to proposition 3 that is otherwise unimportant. We relegate this corollary to Appendix C and show there how it can be applied to construct each of G and H .

Demonstration 4: Modulating the Flicker Frequency of Spatial Noise with a Drifting Sinusoid

We shall construct the random stimulus H of N frame blocks indexed $0, 1, \dots, N - 1$, each composed of M rectangles indexed $0, 1, \dots, M - 1$ from left to right. Let $\rho_0, \rho_1, \dots, \rho_{M-1}$ be pairwise independent random variables, each uniformly distributed on $[-\pi, \pi)$. Let C be a contrast value. For all $(x, y, t) \in \mathbf{Z}^3$, set

$$H[x, y, t] = C \cos \left(4\pi \left(1 + \cos \left(2\pi \left(\frac{m}{M} - \frac{n}{N} \right) \right) \right) + \rho_m \right)$$

for m indexing the rectangle containing (x, y) and n indexing the frame block containing t . The demonstration that H is drift balanced is given in Appendix C.

A realization of H , with $N = 32$ and $M = 128$, is shown in Fig. 4d. In frame block 0, the rectangles are assigned random contrasts between C and $-C$ (as a consequence of their independent, random phases). Thereafter, for $m = 0, 1, \dots, M - 1$, the contrast of the m th rectangle is modulated by a cosine whose phase is itself a sinusoidal function of the rectangle and the frame block. Since, however, a sinusoid's frequency is the derivative of its phase (and since the derivative of a sinusoid is a sinusoid of the same frequency), we observe that H modulates, with a drifting sinusoid, the frequency of (spatially random-phased) sinusoidal flicker.

In demonstration 4 the contrast oscillation rate of each rectangle speeds up and slows down sinusoidally throughout the presentation. Regions of equal oscillation rate (crests of rapid sinusoidal flicker separated by troughs of slow modulation) sweep at a constant rate from left to right across the viewing field.

Method

The conditions under which H was presented to subjects CC and DY were the same as those governing the display of K (of demonstration 2). Each frame block lasted 1/60 sec, each spatial rectangle measured 2 deg (vertical) \times 1/8 deg (horizontal), and the contrast $C = 0.25$.

Results

Interestingly, despite the striking diagonal contours marking the (x, y) pattern of Fig. 4d, both subjects reported that the motion of H was generally more ambiguous than those of the other stimuli. CC (DY) reported apparent motion in the drift direction of the sinusoid modulating frequency of contrast oscillation on 28 (23) of 30 trials.

Demonstration 5: Modulating the Contrast of Flickering Noise with a Drifting Sinusoid

The random stimulus G is made up of N frame blocks indexed $0, 1, \dots, N - 1$, each containing M rectangles indexed $0, 1, \dots, M - 1$ from left to right. Let $\rho_0, \rho_1, \dots, \rho_{M-1}$ be pairwise independent random variables, each uniformly distributed on $[-\pi, \pi)$. Let C be some contrast value; then, for any $(x, y, t) \in \mathbf{Z}^3$, set

$$G[x, y, t] = \frac{C}{2} \left(\cos \left(2\pi \left(\alpha \frac{m}{M} - \beta \frac{n}{N} \right) \right) + 1 \right) \times \cos \left(2\pi \gamma \frac{n}{N} + \rho_m \right),$$

where m indexes the rectangle containing (x, y) and n index-

es the frame block containing t . The proof that G is drift balanced is given in Appendix C.

A realization of G with $M = 128$, $N = 32$, $\alpha = \beta = 2$, and $\gamma = 3$ is shown in Fig. 4e. As does K of demonstration 2, G generates its apparent motion by modulating contrast as a drifting sinusoidal function of the rectangle and the frame block. However, whereas the background whose contrast is being modulated in K is a static row of rectangles randomly painted C or $-C$, the background whose power is modulated in G is a row of rectangles sinusoidally flickering between C and $-C$; each rectangle m has a randomly assigned phase (ρ_m) and is flickering at the rate of $3/32$ cycles/frame block (as a consequence of the term $2\pi 3n/32$).

The contrast of G 's flickering rectangle row is modulated by the factor

$$\cos\left(2\pi\left(\frac{2m}{128} - \frac{2n}{32}\right)\right) + 1,$$

which sweeps peaks (two per frame) of high-contrast flicker separated by troughs of mean gray across the viewing field from left to right.

Method

The conditions governing the display of G were the same as those for K (and H): Frame blocks lasted 1/60 sec, spatial rectangles measured 2 deg (vertical) \times 1/8 deg (horizontal), and $C = 0.25$.

Results

CC (DY) registered apparent motion in the drift direction of the sinusoid modulating noise contrast in G on 30 (26) of 30 trials.

Conclusions

In this section we have demonstrated five drift-balanced random stimuli whose apparent motion is perceived in one consistent direction in more than 90% of trials by two observers. Indeed, many other observers have viewed these stimuli, and no one has yet failed to perceive their consistent motion. As is discussed in Section 8 below, these stimuli are microbalanced in addition to being drift balanced; that is, they remain drift balanced after windowing by arbitrary space-time-separable functions. We conclude that there is a large class of random stimuli whose apparent motion contradicts the MFFC principle of motion perception.

There are many kinds of drift-balanced and microbalanced random stimuli that were not represented among the demonstrations described here. In this paper we have restricted ourselves to stimuli that assign constant values in the vertical dimension of space. Dropping this constraint opens the door to a broad range of other drift-balanced and microbalanced random stimuli. In particular, a large class of displays that yield apparent motion is generated by defining two spatiotemporal texture fields, A and B, at each point $(x, y, t) \in \mathbf{Z}^3$ and moving a boundary that admits light only from field A on one side and only from B on the other. Many instances of this kind of apparent motion, including those proposed by Victor,¹⁸ can easily be shown to be microbalanced.¹⁹

7. REICHARDT-DETECTOR CHARACTERIZATION OF DRIFT-BALANCED RANDOM STIMULI

A point-delay Reichardt detector is a simple device that was proposed originally by Reichardt²⁰ to explain the vision of beetles. Its basic principle, the autocorrelation of inputs at nearby visual locations, underlies most of the currently predominant models of human motion perception. We define the Reichardt detector in terms of two subunits, designated for convenience as the left and right half-detectors. Both half-detectors are defined with respect to the same two (spatial) locations (x, y) and (p, q) in \mathbf{Z}^2 and for some fixed nonnegative number δ_t of frames. These oppositely oriented detectors are pitted additively against each other. A left half-detector r_{left} [implicitly indexed by (x, y) , (p, q) , and δ_t] computes the covariance over time of the contrast at point (x, y) at time t with the contrast at point (p, q) at time $t - \delta_t$ throughout the display of an arbitrary stimulus I . For r_{right} , t and $t - \delta_t$ are reversed. The computation performed by r is given by

$$r(I) = r_{\text{left}}(I) - r_{\text{right}}(I) = \sum_{t \in \mathbf{Z}} I[x, y, t]I[p, q, t - \delta_t] - \sum_{t \in \mathbf{Z}} I[x, y, t - \delta_t]I[p, q, t].$$

When $r(I) < 0$, it indicates motion from (x, y) to (p, q) .

Figure 5 illustrates a block-diagram representation of the Reichardt half-detectors and the Reichardt full detector. The box containing (x, y) [respectively, (p, q)] is a contrast gauge, inputting the contrast at point (x, y) [(p, q)] for each successive frame t . Each of the boxes containing δ_t is a delay filter. At frame t , each delay box outputs the value entered into it at frame $t - \delta_t$. Each of the boxes marked with an \times outputs the product of its two inputs at any frame t . Each of the boxes marked with a Σ accumulates the output from the multipliers over all the frames. Finally, the box marked with a $-$ outputs the difference of its inputs at any frame t .

To see how the detector shown in Fig. 5c works, consider a point of light moving across a dark visual field so as to cross first (x, y) and then (p, q) . If the spot is moving at the proper rate [so that it starts crossing (p, q) after precisely δ_t frames], then the output from the right-hand multiplier will be high as the dot passes over (p, q) . In contrast, the output from the left-hand multiplier will be low throughout the presentation of the moving dot, since, at any frame, at least one of its input channels is contributing a value near zero. Thus the output of the detector is negative. On the other hand, if the dot passes first over (p, q) and then over (x, y) , the detector's response is positive. In this simple case, the sign of the detector's output does a good job of signaling the direction of the dot's motion.

However, the point-delay Reichardt detector is highly vulnerable to aliasing. Imagine a train of evenly spaced dots passing at some speed s first over (x, y) and then over (p, q) . For any s , it is easy to adjust the spacing between dots so that the output of the Reichardt detector of Fig. 5c signals rightward motion, leftward motion, or no motion at all.

Despite the shortcomings of the simple Reichardt detector, there is something appealing about its fundamental au-

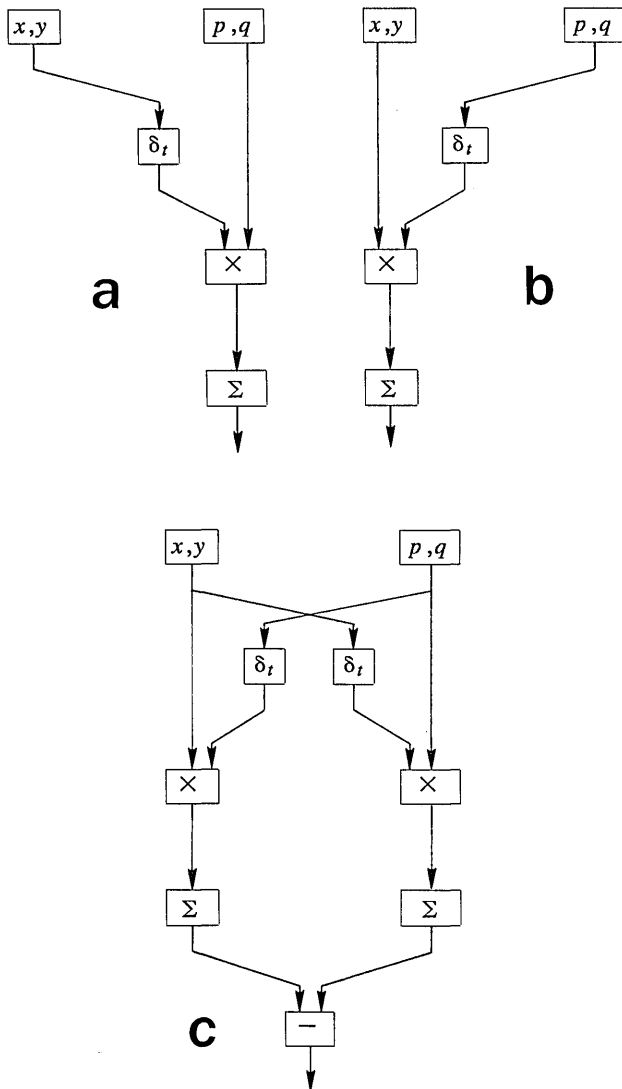


Fig. 5. Point-delay Reichardt detector and its component half-detectors. a, The right half-detector computes the covariance of the contrast fluctuations of the input stimulus at point (p, q) with the fluctuations δ_t frames earlier at point (x, y) : (x, y) and (p, q) register signal contrast frame by frame. The contrast of the current frame at pixel (p, q) is multiplied by the contrast at pixel (x, y) δ_t frames in the past. (The box labeled δ_t outputs the input it received δ_t frames ago.) The output from the multiplier is accumulated over all the frames of the display. b, In a similar fashion, the left half-detector computes the covariance of the contrast fluctuations of the input stimulus at point (x, y) with the fluctuations δ_t frames earlier at point (p, q) . c, The full point-delay Reichardt detector outputs the difference between the left and right half-detectors. A positive response thus signals leftward motion; a negative response signals rightward motion.

tocorrelation principle. Various elaborations of Reichardt models were developed and studied in detail by van Santen and Sperling,^{1,2,21} who proved that the apparently different models of Adelson and Bergen³ and Watson and Ahumada⁴ were essentially special types of elaborated Reichardt detectors (ERD's). All these models retain the basic delay-and-compare structure of the simple detector diagrammed in Fig. 5c. However, this simple detector is generalized in the following ways: (i) the point detectors at (x, y) and (p, q) are replaced by spatial receptive fields (that is, each receptive

field applies an array of weights to the stimulus impinging upon its region of the retina, and it outputs the sum of the weighted contrast values), (ii) the temporal point delays before the multipliers are replaced by temporal filters, and (iii) the temporal accumulators after the multipliers are replaced by temporal filters. Van Santen and Sperling² showed that further additions (e.g., more temporal filters added here and there) do not augment the capabilities of this ERD.

It was widely assumed that, ideally, a good motion detector should behave as a frequency-domain power analyzer.^{1-6,21-23} (This is the assumption called into question by the demonstration of good apparent motion in drift-balanced stimuli.) The simple point-delay Reichardt detector falls short of this ideal: it is not a good Fourier analyzer. The various elaborations of Reichardt detectors can be viewed as attempts to improve their performance as frequency-domain power analyzers.

There is another way to use the Reichardt mechanism as the basis of a motion-perception model. Indeed, as we shall observe, it is possible to build a perfect Fourier power analyzer by using only the simplest point-delay half-detectors.

Our main purpose in this section, however, is to provide an alternative characterization of the class of drift-balanced random stimuli, in terms of the expected responses of point-delay Reichardt detectors to members of this class. We prove the following proposition: For any integers δ_x , δ_y , and δ_t , form the class $C_{\delta_x, \delta_y, \delta_t}$ of all point-delay Reichardt detectors conforming to Fig. 5c [with (x, y) and (p, q) ranging throughout \mathbb{Z}^2] such that $(x, y) - (p, q) = (\delta_x, \delta_y)$, and call $C_{\delta_x, \delta_y, \delta_t}$ trivial if either $(\delta_x, \delta_y) = (0, 0)$ or $\delta_t = 0$; that is, $C_{\delta_x, \delta_y, \delta_t}$ is trivial if its member detectors fail to separate, either in space or time, the points whose contrast they compare. I is then drift balanced iff the expected pooled response of every nontrivial class of point-delay Reichardt detectors is 0. We now proceed more formally.

Definition 4: Autocorrelation

Let I be a random stimulus. Then for any $\delta = (\delta_x, \delta_y, \delta_t) \in \mathbb{Z}^3$, define the autocorrelation, \mathbf{H}_I , by

$$\mathbf{H}_I[\delta_x, \delta_y, \delta_t] = \sum I[x, y, t]I[p, q, r],$$

where the sum is taken over all pairs $(x, y, t), (p, q, r) \in \mathbb{Z}^3$ for which $(x, y, t) - (p, q, r) = (\delta_x, \delta_y, \delta_t)$. Define the full-detector pooler, \mathbf{R}_I , by setting

$$\mathbf{R}_I[\delta_x, \delta_y, \delta_t] = \mathbf{H}_I[\delta_x, \delta_y, \delta_t] - \mathbf{H}_I[-\delta_x, -\delta_y, \delta_t].$$

We use \mathbf{H}_I to denote the autocorrelation of I because, for any $(\delta_x, \delta_y, \delta_t)$, $\mathbf{H}_I[\delta_x, \delta_y, \delta_t]$ collects the sum of the responses to I of all the half-detectors conforming to Fig. 5b, with δ_t delay filters, such that $(x, y) - (p, q) = (\delta_x, \delta_y)$. The half-detectors corresponding to Fig. 5a are pooled by $\mathbf{H}_I[-\delta_x, -\delta_y, \delta_t]$. Thus $\mathbf{R}_I[\delta_x, \delta_y, \delta_t]$ pools the output of all full Reichardt detectors corresponding to Fig. 5c, with $(x, y) - (p, q) = (\delta_x, \delta_y)$ (and δ_t delay filters).

Observation 2

For any random (or nonrandom) stimulus I and any $\delta = (\delta_x, \delta_y, \delta_t) \in \mathbb{Z}^3$,

$$\mathbf{H}_I[\delta] = \mathbf{H}_I[-\delta].$$

The proof is trivial.

In order to reclaim Fourier motion information from the half-detector output, note first that, for any random stimulus I ,

$$|\bar{I}(\omega, \theta, \tau)|^2 = \sum I[x, y, t]I[p, q, r] \times \exp(j(\omega(x-p) + \theta(y-q) + \tau(t-r))), \quad (4)$$

where the sum is taken over all $(x, y, t), (p, q, r) \in \mathbf{Z}^3$. We can now collect terms of the sum in Eq. (4) that have identical exponential factors to obtain

$$|\bar{I}(\omega, \theta, \tau)|^2 = \sum \mathbf{H}_I[\delta_x, \delta_y, \delta_t] \exp(j(\omega\delta_x + \theta\delta_y + \tau\delta_t)), \quad (5)$$

where the sum is over all $(\delta_x, \delta_y, \delta_t) \in \mathbf{Z}^3$.

Equation (5) shows that point-delay half-detectors, by themselves, contain all the information about the distribution of I 's power in the Fourier domain (because \mathbf{H}_I depends on only the output of half-detectors to I).

The next definition is useful for proving the main result of this section.

Definition 5: Power Difference between Oppositely Drifting Fourier Components

For any random stimulus I and any $\omega, \theta, \tau \in \mathbb{R}$, set

$$\Delta_I(\omega, \theta, \tau) = |\bar{I}(\omega, \theta, \tau)|^2 - |\bar{I}(\omega, \theta, -\tau)|^2.$$

Note that any random stimulus I is drift balanced iff $E[\Delta_I(\omega, \theta, \tau)] = 0$ for all $\omega, \theta, \tau \in [0, 2\pi)$. Some facts about Δ_I are worth noting. First,

$$\begin{aligned} \Delta_I(\omega, \theta, \tau) &= \sum (\mathbf{H}_I[\delta_x, \delta_y, \delta_t] - \mathbf{H}_I[\delta_x, \delta_y, -\delta_t]) \times \exp(j(\omega\delta_x + \theta\delta_y + \tau\delta_t)) \\ &= \sum (\mathbf{H}_I[\delta_x, \delta_y, \delta_t] - \mathbf{H}_I[-\delta_x, -\delta_y, \delta_t]) \times \exp(j(\omega\delta_x + \theta\delta_y + \tau\delta_t)) \\ &= \sum \mathbf{R}_I[\delta_x, \delta_y, \delta_t] \exp(j(\omega\delta_x + \theta\delta_y + \tau\delta_t)), \end{aligned}$$

where each sum is over all $(\delta_x, \delta_y, \delta_t) \in \mathbf{Z}^3$. The first identity depends on the fact that

$$\begin{aligned} |\bar{I}(\omega, \theta, -\tau)|^2 &= \sum \mathbf{H}_I[\delta_x, \delta_y, \delta_t] \exp(j(\omega\delta_x + \theta\delta_y - \tau\delta_t)) \\ &= \sum \mathbf{H}_I[\delta_x, \delta_y, -\delta_t] \exp(j(\omega\delta_x + \theta\delta_y + \tau\delta_t)). \end{aligned}$$

The second identity follows from observation 2.

Next note that any term

$$(\mathbf{H}_I[\delta_x, \delta_y, \delta_t] - \mathbf{H}_I[\delta_x, \delta_y, -\delta_t]) \exp(j(\omega\delta_x + \theta\delta_y + \tau\delta_t))$$

in the sum yielding $\Delta_I(\omega, \theta, \tau)$ is obviously 0 if $\delta_t = 0$. On the other hand, this term is equal (by observation 2) to

$$(\mathbf{H}_I[\delta_x, \delta_y, \delta_t] - \mathbf{H}_I[-\delta_x, -\delta_y, \delta_t]) \exp(j(\omega\delta_x + \theta\delta_y + \tau\delta_t)),$$

which is evidently 0 if $\delta_x = \delta_y = 0$. This goes to show that for any $\delta_x, \delta_y, \delta_t \in \mathbf{Z}$, any class of Reichardt half-detectors, each of whose members has (i) no separation between spatial receptors or (ii) a delay factor of 0, does not influence $\Delta_I(\omega, \theta, \tau)$.

The following lemma summarizes these observations.

Lemma 2

For any random stimulus I , any $\omega, \theta, \tau \in \mathbb{R}$,

$$\Delta_I(\omega, \theta, \tau) = \sum \mathbf{R}_I[\delta_x, \delta_y, \delta_t] \exp(j(\omega\delta_x + \theta\delta_y + \tau\delta_t)), \quad (6)$$

where the sum is taken over all integers δ_x, δ_y , and δ_t such that $\delta_t \neq 0$ and either $\delta_x \neq 0$ or $\delta_y \neq 0$.

Obviously, if

$$E[\mathbf{R}_I[\delta_x, \delta_y, \delta_t]] = 0$$

for all δ_x, δ_y , and δ_t indexing the sum in Eq. (6), then $\Delta_I(\omega, \theta, \tau) = 0$. This proves half of the following proposition.

Proposition 4

A random stimulus is drift balanced iff the expected pooled output from every nontrivial class of Reichardt detectors is 0; that is, any random stimulus I is drift-balanced iff

$$E[\mathbf{R}_I[\delta_x, \delta_y, \delta_t]] = 0 \quad (7)$$

for all integers δ_x, δ_y , and δ_t such that $\delta_t \neq 0$ and $(\delta_x, \delta_y) \neq (0, 0)$.

Proof

We have already observed that Eq. (7) implies that I is drift balanced. It remains to be proved that Eq. (7) holds whenever I is drift balanced. Accordingly, let Q be the set of all $(\delta_x, \delta_y, \delta_t)$ for which $\delta_t \neq 0$ and $(\delta_x, \delta_y) \neq (0, 0)$, and suppose that, for any $\omega, \theta, \tau \in [0, 2\pi)$,

$$E[\Delta_I(\omega, \theta, \tau)] = 0.$$

When we take expectations of both sides of Eq. (6), and multiply each side of the resulting identity by its conjugate, we obtain

$$\begin{aligned} E^2[\Delta_I(\omega, \theta, \tau)] &= \sum E[\mathbf{R}_I[\delta_x, \delta_y, \delta_t]] E[\mathbf{R}_I[\delta_p, \delta_q, \delta_r]] \\ &\quad \times \exp(j(\omega(\delta_x - \delta_p) + \theta(\delta_y - \delta_q) + \tau(\delta_t - \delta_r))), \quad (8) \end{aligned}$$

where the sum is over all $(\delta_x, \delta_y, \delta_t), (\delta_p, \delta_q, \delta_r) \in Q$. However, recalling that

$$\begin{aligned} &\int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \exp(j(\omega(\delta_x - \delta_p) + \theta(\delta_y - \delta_q) + \tau(\delta_t - \delta_r))) d\omega d\theta d\tau \\ &= \int_0^{2\pi} \exp(j\omega(\delta_x - \delta_p)) d\omega \int_0^{2\pi} \exp(j\theta(\delta_y - \delta_q)) d\theta \\ &\quad \times \int_0^{2\pi} \exp(j\tau(\delta_t - \delta_r)) d\tau \\ &= \begin{cases} (2\pi)^3 & \text{if } \delta_x = \delta_p, \quad \delta_y = \delta_q, \quad \delta_t = \delta_r, \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

we find that when we integrate both sides of Eq. (8) over the interval $[0, 2\pi)^3$ and divide through by $(2\pi)^3$, we obtain

$$\sum E^2[\mathbf{R}_I[\delta_x, \delta_y, \delta_t]] = \frac{1}{8\pi^3} \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} E^2[\Delta_I(\omega, \theta, \tau)] d\omega d\theta d\tau.$$

where the sum is over all $(\delta_x, \delta_y, \delta_t) \in Q$. But the right-hand side of this identity is 0 by assumption. Thus, since each term in the left-hand sum is nonnegative, each must be 0. ■

For current purposes, the importance of the Reichardt-detector characterization of the class of drift-balanced random stimuli (established in proposition 4) is that it provides

easy access to the principal results concerning the critical subclass of drift-balanced random stimuli that we call microbalanced. This is the focus of Section 8.

8. MICROBALANCED RANDOM STIMULI

Consider the following two-frame-block stimulus S : In frame block 0, a bright spot (call it spot 0) appears. In frame block 1, spot 0 disappears, and two new spots appear, one on each side of spot 0. On the one hand, it is clear (from proposition 4) that S is drift balanced. On the other hand, it is equally clear that a Fourier-based motion detector whose spatial reach encompassed the location of spot 0 and only one of the flashes in frame block 1 might be stimulated strongly in a fixed direction by S . Although S is drift balanced, some local Fourier motion detectors would be stimulated strongly and systematically by S . These detectors can be selected differentially by spatial windowing, and thereby a drift-balanced stimulus S can be converted into a non-drift-balanced stimulus.

In this section we introduce the class of microbalanced random stimuli, a subclass of drift-balanced random stimuli, any member I of which is guaranteed not to stimulate Fourier-power motion detectors in any systematic way, regardless of any space-time-separable window interposed between I and the detector. As we shall prove in proposition 8 below, I possesses this property if I satisfies the following definition.

Definition 6: Microbalanced Stimulus

Call any random stimulus I *microbalanced* iff, for any $(x, y, t), (x', y', t') \in \mathbf{Z}^3$,

$$E[I[x, y, t]I[x', y', t']] = E[I[x, y, t']I[x', y', t]].$$

Obviously, for any random spatial function f and temporal random function g ,

$$E[f[x, y]g[t]f[x', y']g[t']] = E[f[x, y]g[t']f[x', y']g[t]],$$

yielding the following proposition.

Proposition 5

Any space-time-separable random stimulus is microbalanced.

A related result is stated in the next proposition.

Proposition 6

Any invariant microbalanced stimulus I is space-time-separable.

Proof

If $I = \mathbf{0}$, there is nothing to prove (since, obviously, $\mathbf{0}$ is space-time separable). Otherwise we choose a point $(x', y', t') \in \mathbf{Z}^3$, for which $I[x', y', t'] \neq 0$, and, for all $(x, y, t) \in \mathbf{Z}^3$, we define

$$f(x, y) = I[x, y, t']$$

and

$$g(t) = \frac{I[x', y', t]}{I[x', y', t']}. .$$

If either $(x, y) = (x', y')$ or $t = t'$, then immediately we obtain

$$I[x, y, t] = f(x, y)g(t).$$

On the other hand, if $(x, y) \neq (x', y')$ and $t \neq t'$, I 's invariance and microbalancedness together imply that

$$I[x, y, t] = \frac{I[x, y, t']I[x', y', t]}{I[x', y', t']} = f(x, y)g(t). \quad \blacksquare$$

An important property of microbalanced random stimuli that sets them apart from the more general class of drift-balanced random stimuli is explained in proposition 7.

Proposition 7

The product of independent microbalanced random stimuli I and J is microbalanced.

Proof

For any $(x, y, t), (x', y', t') \in \mathbf{Z}^3$,

$$\begin{aligned} E[IJ[x, y, t]IJ[x', y', t']] &= E[I[x, y, t]I[x', y', t']]E[J[x, y, t]J[x', y', t']] \\ &= E[I[x, y, t']I[x', y', t']]E[J[x, y, t']J[x', y', t']] \\ &= E[IJ[x, y, t']IJ[x', y', t']]. \quad \blacksquare \end{aligned}$$

Earlier in this section we showed, by using the example of a single spot splitting into two adjacent spots, that a drift-balanced random stimulus (S) can systematically stimulate motion detectors that operate on restricted regions of S . With proposition 8 we shall establish that all and only those random stimuli that are microbalanced avoid the systematic stimulation of all local (and global) Fourier-power detectors. The following lemma eases the proof of this important fact.

Lemma 3

Any microbalanced random stimulus is drift balanced.

Proof

Let I be microbalanced. From proposition 4, I is drift balanced iff $E[\mathbf{H}_I[\delta_x, \delta_y, \delta_t]] = E[\mathbf{H}_I[\delta_x, \delta_y, -\delta_t]]$ for any offset $(\delta_x, \delta_y, \delta_t) \in \mathbf{Z}^3$, such that $(\delta_x, \delta_y) \neq (0, 0)$ and $\delta_t \neq 0$. However, since I is microbalanced, we note that for any such $(\delta_x, \delta_y, \delta_t)$,

$$\begin{aligned} E[\mathbf{H}_I[\delta_x, \delta_y, \delta_t]] &= E[\sum I[x, y, t]I[x - \delta_x, y - \delta_y, t - \delta_t]] \\ &= \sum E[I[x, y, t]I[x - \delta_x, y - \delta_y, t - \delta_t]] \\ &= \sum E[I[x, y, t - \delta_t]I[x - \delta_x, y - \delta_y, t]] \\ &= E[\sum I[x, y, t - \delta_t]I[x - \delta_x, y - \delta_y, t]] \\ &= E[\sum I[x, y, t]I[x - \delta_x, y - \delta_y, t + \delta_t]] \\ &= E[\mathbf{H}_I[\delta_x, \delta_y, -\delta_t]], \end{aligned}$$

where each of the sums is over all $(x, y, t) \in \mathbf{Z}^3$. \blacksquare

We can now state the main result of this section.

Proposition 8

For any random stimulus I , the following conditions are equivalent:

- I. I is microbalanced.
- II. For any space-time-separable function W , WI is drift balanced.

Proof

First we prove that condition I implies condition II. Assume that I is microbalanced. By proposition 5, W is also microbalanced; it thus follows proposition 7 that WI is microbalanced and hence drift balanced (from lemma 3).

Next we prove that not condition I implies not condition II. Suppose that I is not microbalanced; then, for some $(x, y, t), (x', y', t') \in \mathbf{Z}^3$,

$$E[I[x, y, t]I[x', y', t']] \neq E[I[x, y, t']I[x', y', t]].$$

[Note that this inequality implies that $(x, y) \neq (x', y')$ and $t \neq t'$.] Let f assign 1 to (x, y) and (x', y') , and let it assign 0 to all other points of \mathbf{Z}^2 ; and let g assign 1 to t and t' and 0 to all other points of \mathbf{Z} . Then the function fgI is zero everywhere except at the points $(x, y, t), (x, y, t'), (x', y', t')$, and (x', y', t') . It is obvious, from proposition 4, that fgI is not drift balanced. In particular,

$$\begin{aligned} E[\mathbf{H}_{fgI}[x - x', y - y', t - t']] &= E[I[x, y, t]I[x', y', t']] \\ &\neq E[I[x, y, t']I[x', y', t']] \\ &= E[\mathbf{H}_{fgI}[x - x', y - y', -(t - t')]]. \blacksquare \end{aligned}$$

The results stated thus far in this section would not be interesting if there were no microbalanced random stimuli that displayed consistent apparent motion. The following result makes it clear that, in fact, all the examples of drift-balanced random stimuli that we considered previously are microbalanced.

Proposition 9

Let Γ be a family of pairwise independent, microbalanced random stimuli, all but at most one of which have an expectation of 0; then any linear combination of Γ is microbalanced.

Proof

Since a microbalanced random stimulus multiplied by a constant remains microbalanced, we assume that the linear combination is a sum; then, for any $(x, y, t), (x', y', t') \in \mathbf{Z}^3$,

$$\begin{aligned} E\left[\sum_{I \in \Gamma} I[x, y, t] \sum_{J \in \Gamma} J[x', y', t']\right] &= \sum_{I \in \Gamma} \sum_{J \in \Gamma} E[I[x, y, t]J[x', y', t']]. \quad (9) \end{aligned}$$

However, whenever $I \neq J$,

$$E[I[x, y, t]J[x', y', t']] = E[I[x, y, t]]E[J[x', y', t']] = 0.$$

Thus Eq. (9) becomes

$$\begin{aligned} \sum_{I \in \Gamma} E[I[x, y, t]I[x', y', t']] &= \sum_{I \in \Gamma} E[I[x, y, t']I[x', y', t']] \\ &= E\left[\sum_{I \in \Gamma} I[x, y, t'] \sum_{J \in \Gamma} J[x', y', t']\right]. \blacksquare \end{aligned}$$

Next we secure the analog of proposition 2.

Proposition 10

The (spatiotemporal) convolution of two independent microbalanced random stimuli is microbalanced.

Proof

It is convenient to write

$$\sum_{a_1, a_2, \dots, a_n}$$

for a sum in which each of the variables a_i ranges over all integers. For any independent random stimuli I and J and any $(x, y, t), (x', y', t') \in \mathbf{Z}^3$,

$$\begin{aligned} E[I * J[x, y, t]I * J[x', y', t']] &= E\left[\sum_{p, q, r} I[x - p, y - q, t - r]J[p, q, r] \right. \\ &\quad \left. \times \sum_{p', q', r'} I[x' - p', y' - q', t' - r']J[p', q', r']\right] \\ &= \sum_{p, q, r, p', q', r'} E[I[x - p, y - q, t - r]I[x' - p', y' - q', t' - r']] \\ &\quad \times E[J[p, q, r]J[p', q', r']]. \end{aligned}$$

But if, in addition, I and J are microbalanced, then this last sum is equal to

$$\begin{aligned} \sum_{p, q, r, p', q', r'} E[I[x - p, y - q, t' - r']I[x' - p', y' - q', t - r]] &\times E[J[p, q, r']J[p', q', r]] \\ &= E\left[\sum_{p, q, r'} I[x - p, y - q, t' - r']J[p, q, r'] \right. \\ &\quad \left. \times \sum_{p', q', r} I[x' - p', y' - q', t - r]J[p', q', r]\right] \\ &= E[I * J[x, y, t']I * J[x', y', t]]. \blacksquare \end{aligned}$$

Response of Reichardt Detectors to Microbalanced Random Stimuli

Two Fourier-analytic motion detectors proposed for psychophysical data^{3,4} can be recast as variants of an ERD.^{2,3} The ERD has many useful properties as a motion detector without regard to its specific instantiation.^{1,2,21}

Figure 6 shows a diagram of the ERD. It consists of spatial receptors characterized by spatial functions f_1 and f_2 , temporal filters g_1^* and g_2^* , multipliers, an adder, and another temporal filter h^* . The spatial receptors f_i ($i = 1, 2$) act on the input stimulus I to produce intermediate outputs,

$$y_i[t] = \sum_{(x, y) \in \mathbf{Z}^2} f_i[x, y]I[x, y, t].$$

At the next stage, each temporal filter g_j^* transforms its input y_j ($i, j = 1, 2$), yielding four temporal output functions: $g_j * y_i$. The left and right multipliers then compute

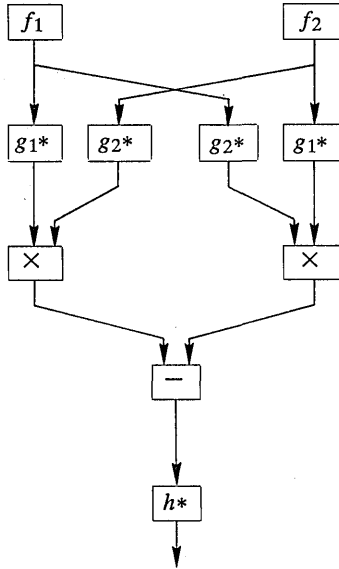


Fig. 6. Diagram of the ERD. Let I be a random stimulus; then, in response to I , for $i = 1, 2$, the box containing the spatial function $f_i: \mathbb{Z}^2 \rightarrow \mathbb{R}$ outputs the temporal function $\sum_{(x,y) \in \mathbb{Z}^2} f_i[x, y]I[x, y, t]$; each of the boxes marked g_i^* outputs the convolution of its input with the temporal function $g_i: \mathbb{Z} \rightarrow \mathbb{R}$; each of the boxes marked with a \times outputs the product of its inputs; the box marked with a $-$ outputs its left input minus its right; and the box containing h^* outputs the convolution of its input with the temporal function $h: \mathbb{Z} \rightarrow \mathbb{R}$.

$$[g_1 * y_1[t]][g_2 * y_2[t]], \quad [g_1 * y_2[t]][g_2 * y_1[t]],$$

respectively, and the differencer subtracts the output of the right multiplier from that of the left multiplier:

$$D[t] = [g_1 * y_1[t]][g_2 * y_2[t]] - [g_1 * y_2[t]][g_2 * y_1[t]].$$

The final output is produced by applying the filter h^* , whose purpose is to appropriately smooth the time-varying differencer output D .

In the following discussion, we write

$$\sum_{a_1, a_2, \dots, a_n}$$

for a sum in which each of the variables a_i ranges over all integers. Given a random stimulus I as the input to the ERD, the output of the differencing component at time B is

$$D[B] = \left[\sum_u g_1[u] \sum_{x,y} f_1[x, y]I[x, y, B - u] \right] \times \left[\sum_t g_2[t] \sum_{p,q} f_2[p, q]I[p, q, B - t] \right] - \left[\sum_u g_1[u] \sum_{p,q} f_2[p, q]I[p, q, B - u] \right] \times \left[\sum_t g_2[t] \sum_{x,y} f_1[x, y]I[x, y, B - t] \right],$$

which can be rewritten as

$$D[B] = \sum_{t,u,p,q,x,y} g_1[u]g_2[t]f_1[x, y]f_2[p, q] \times [I[x, y, B - u]I[p, q, B - t] - I[x, y, B - t]I[p, q, B - u]].$$

However, if I is microbalanced, then (by definition 6) the expectation of the square-bracketed difference is 0, and hence $E[D[B]] = 0$ for any $B \in \mathbb{Z}$, implying the following proposition.

Proposition 11

The expected response of any elaborated Reichardt detector to any microbalanced random stimulus is 0 at every instant in time.

Microbalanced random stimuli, then, compose a subclass of drift-balanced random stimuli with special importance for the investigation of non-Fourier motion perception. In general, the fact that a random stimulus I is drift balanced does not entail that all local areas of I be drift balanced; that is, the window over which the Fourier power analysis of I is carried out is critical to the drift-balancedness of I . This constraint is escaped by microbalanced random stimuli (as a consequence of proposition 8): a random stimulus I is microbalanced iff, for any space-time-separable function W , the random stimulus WI (the result of windowing I by W) is drift balanced.

9. RECOVERY OF MOTION FROM MICROBALANCED RANDOM STIMULI

Nonlinear Transformations Hypothesis

The most plausible explanation for the recovery of motion from drift-balanced random stimuli posits one or more nonlinear transformations that are routinely applied to the visual input signal to generate a new signal, which is then subjected to ordinary frequency-domain power analysis.

Consider, for instance, random stimuli such as those described in demonstrations 1, 2, and 5 (Figs. 4a, 4b, and 4e), whose motion depends on spatiotemporal modulation of noise contrast. For concreteness, we focus on I , the contrast-reversing bar of demonstration 1 (Fig. 4a). The apparent motion exhibited by I might result from a power analysis in the frequency domain of a rectified version of the original signal: for example, a transformation of the signal I such as R_I, S_I, T_I^+ , or T_I^- , where

- (i) $R_I[x, y, t] = |I[x, y, t]|$ (full-wave rectification),
- (ii) $S_I[x, y, t] = I[x, y, t]^2$ (full-wave power rectification),
- (iii) $T_I^+[x, y, t] = \max\{I[x, y, t], 0\}$ (positive half-wave rectification),
- (iv) $T_I^-[x, y, t] = \min\{I[x, y, t], 0\}$ (negative half-wave rectification).

R_I and S_I both transform I into a rectangle moving in a series of brief steps from left to right, while T_I^+ and T_I^- map I into a similar such moving rectangle, which randomly disappears and reappears in the course of its left-to-right traversal. The MFFC principle applied to any of these transformations of I would indicate motion to the right (see Fig. 7a). In the

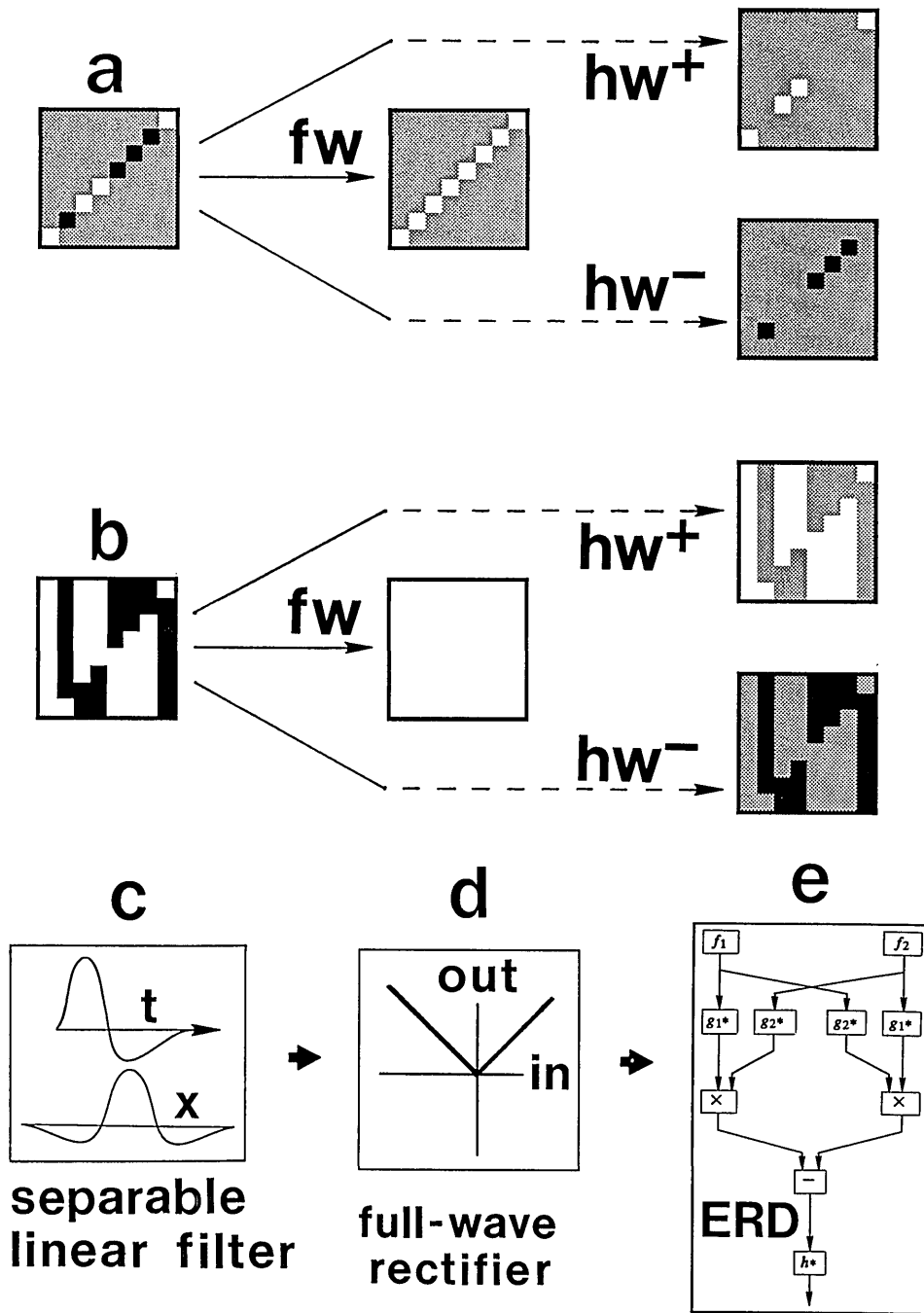


Fig. 7. Consequences of full-wave and half-wave rectification. a, Space-time representation of a traveling, contrast-reversing bar; full-wave (fw) rectified representation; and positive (hw^+) and negative (hw^-) half-wave rectified representations, showing that either of these rectifications suffices to expose the motion to Fourier motion-energy analysis. b, Space-time representation of a traveling contrast reversal of a random bar pattern; full-wave (fw) rectified representation; positive (hw^+) and negative (hw^-) half-wave rectified representations, showing that none of these rectifications exposes motion. The analysis system for second-order motion stimuli is shown in the bottom row: c, the signal is linearly filtered (the impulse response of an appropriate space-time-separable linear filter is shown); d, the filtered signal is full-wave rectified; and e, it is subjected to motion-energy analysis (e.g., by an ERD). This is a sufficient sequence of operations to expose the directional motion in all the demonstrations of this paper.

realm of spatial visual perception, rectification transformations were proposed by various authors to mediate boundary formation and texture segregation.²⁴⁻²⁸ Logarithmic intensity compression was also proposed,²⁹⁻³² because of its physiological plausibility, although it is less effective than rectification.

Although any one of the rectification transformers would expose the motion information buried in I to frequency-domain power analysis, the same is not true of the traveling

contrast-reversal J defined in demonstration 3 (Fig. 4c). Full-wave rectification of J yields a constant output. Half-wave rectification merely yields another drift-balanced random stimulus: $T_{J^+} = (J + 1)/2$ and $T_{J^-} = (1 - J)/2$. These relations are illustrated in Fig. 7b. The motion of J does not emerge directly from any of these forms of rectification.

For the traveling, random contrast-reversal J (demonstration 3, Fig. 4c), a time-dependent linear operator such as temporal differentiation is required to transform it into a

signal from which motion information can be extracted after rectification. (Indeed, the partial derivative of J with respect to time is I .)

Consider the space-time-separable bandpass filtering that is usually assumed to occur in low-level visual processing. If such linear filtering were applied to any of the demonstrations considered in this paper, and if it were followed by any of the rectification operations considered above, it would suffice to expose the motion of any of these demonstrations to Fourier power analysis. Figure 7 illustrates the sequence of filtering, rectifying, and motion-power analysis. A central issue concerning drift-balanced random stimuli thus emerges: given the (largely unexplored) range of drift-balanced random stimuli that elicit apparent motion, what is the simplest array of transformations of the input signal that suffices to expose (to frequency-domain power analysis) the motion information carried by all the various types of drift-balanced random stimuli?

What is the Purpose of Having Detectors for Drift-Balanced Motion?

From a systems point of view, there is a problem in linearly combining the information from many linear sensors (for example, motion-sensitive sensors) because there is nothing gained by the combination that could not have been accomplished by a single, large sensor. For an advantage to be gained from the combination, this information must be nonlinearly related to the input. Nonlinearly computed quantities such as power and information are combined most usefully. In many classical detection problems the ideal detector is a power detector; that is, the power of the component elements is summed to form the decision variable.^{33,34} When it comes to detecting motion, it would be surprising if generally similar considerations did not apply in combining information from various locations of the visual field and from detectors of various sizes. Indeed, the MFFC theories normally use motion detectors that compute Fourier power.¹⁻⁶

Assuming that evolution chooses detection modes because of their advantages, what is surprising about the detection of drift-balanced motion is that the advantages of nonlinear combination are already available at the earliest stages of sensory analysis. Ultimately, to appreciate why this is so requires ecological analysis of the visual world. Obviously, the ecological problem cannot be resolved by armchair speculation. On the other hand, given that combination mechanisms operate with rectified inputs, it is not surprising that the mechanisms that detect drift-balanced motion seem to be of a much larger scale than the Fourier mechanisms.³⁵ A possibly related observation is that the apparent motion in various drift-balanced random stimuli that we have considered here tends to diminish with the retinal eccentricity of the presentation.¹¹ However, it remains to be determined how much of this drop-off of apparent motion should be attributed to the effective decrease in visual spatial sampling rate with retinal eccentricity.

10. UTILITY OF RANDOM STIMULI AS A RESEARCH TOOL

A general advantage of random stimuli compared with repeated stimuli is that the responses to a repeated stimulus might be mediated by any of its features, including artifact-

tual stimulus features that are not anticipated by the experimenter. Responses to random stimuli represent the responses to the properties that distinguish a class of stimuli, and these tend to be more general and more readily specifiable than the properties of a single stimulus. Thus, by generalizing the notion of a stimulus to that of a random stimulus, we obtain a much more extensive and adaptable set of tools for studying perception.

In the study of motion perception, microbalanced random stimuli play a crucial role: they avoid the complications introduced by the spatial windowing that is unavoidably performed by motion-perception units. Avoiding the possible artifacts of windowing is particularly important in interpreting the responses of single visual neurons. Only a microbalanced random stimulus is guaranteed to contain no consistent Fourier components, regardless of how that stimulus may be centered or fail to be centered in a given neuron's receptive field or in the observer's field of view. It is possible for drift-balanced (but not microbalanced) random stimuli to produce systematic Fourier motion components in receptive fields of particular neurons that happen to be placed advantageously with respect to those stimuli. Only microbalanced random stimuli necessarily require non-Fourier operations in order to yield motion perception.

An invariant stimulus is microbalanced (thereby avoiding the windowing problem) only if it is space-time separable (proposition 6). Unfortunately, there are no examples of space-time-separable stimuli that yield a strong, consistent perception of motion. Thus random microbalanced stimuli that yield strong perceived motion offer a unique tool for the investigation of non-Fourier motion perception.

11. NON-FOURIER STIMULUS ANALYSIS IN OTHER SENSORY DOMAINS

Spatial Vision

One-dimensional motion stimuli in (x, t) can be represented as two-dimensional stimuli in (x, y) . From the point of view of systems analysis, the (x, t) and (x, y) representations are equivalent: motion in (x, t) is equivalent to orientation in (x, y) . There are inevitably some physical restrictions that apply in the time domain,² so that x and t cannot be so symmetrical with respect to each other as x and y . For example, in human motion detectors, summation over time (of comparator output) occurs within a single detector; summation over space occurs between detectors.

The space-time asymmetry in motion can be made obvious by adding two gratings. Thus, when a drifting sine-wave grating of frequency (ω_x, ω_t) is added to a stationary sine pattern of frequency $(\omega_x, 0)$ (a standing grating), the apparent motion is normally visible; when it is added to $(0, \omega_t)$ (a uniform, flickering field), the apparent motion may either be normal or be reversed, depending on the phase relations.² In the space domain, both combinations are equivalent.

The fact that all the (x, y) spatial illustrations in the figures of (x, t) motions were visible as oriented textures demonstrates that the same or similar nonlinear dynamics are involved in the extraction of orientation as are involved in the extraction of direction of motion. Indeed, we have yet to discover an (x, t) stimulus that is perceived as moving and that is not perceived as oriented texture in an (x, y)

representation. This suggests that the human array of pattern-analytic detectors is at least as rich as the motion-analytic array.

Audition

Obviously, a one-dimensional signal, such as an auditory signal (which depends only on time), cannot be drift balanced. Nonetheless, certain auditory phenomena bear a resemblance to some of the visual effects that we have been considering.

It has long been recognized that the auditory system analyzes sound-pressure waveforms into their component sinusoidal frequencies and that these frequency components correspond, at least to a first approximation, to the sensation of pitch. Indeed, the cochlea functions largely as a mechanical frequency analyzer. In addition to pure frequency analysis, especially at periodicities below 300 Hz, another mechanism, periodicity analysis, also comes into play. One of the best demonstrations is an experiment by Miller and Taylor.³⁶

Some background facts about this experiment are useful here. A broad-spectrum noise N is a random function of time such that the expected power of all Fourier components in N is equal. It is easy to show that any random function N that assigns pairwise independent random variables, all with mean 0, to distinct points in time is a broad-spectrum noise. Obviously, multiplying any such random function N by an arbitrary nonrandom function f yields yet another broad-spectrum noise, since the values assigned by fN remain pairwise independent, each with mean 0.

In the experiment by Miller and Taylor, listeners heard a broad-spectrum noise that was modulated on and off (multiplied) by a square wave of frequency f . Thus the stimulus generated by Miller and Taylor had a uniform expected power over all temporal frequencies. When f was less than ~ 10 Hz, the perception corresponded to the physical reality of interrupted noise. At frequencies between 40 and 200 Hz, the interrupted noise was perceived to have a pitch that corresponded to the interruption frequency. That observers perceive a pitch implicates some mechanism other than frequency analysis. Whereas a rectifying nonlinearity was not proposed explicitly by Miller and Taylor, it is the obvious intermediate step in periodicity pitch perception.

12. FINAL REMARKS

We have given precise definition to the notion of a random stimulus and focused our attention on the subclasses of drift-balanced and microbalanced random stimuli as being especially interesting for the study of visual perception. We first showed that the (spatiotemporal) convolution of independent drift-balanced random stimuli is drift balanced.

Proposition 3 (which states that the sum of drift-balanced random stimuli is drift balanced when the elements are pairwise independent and all but at most one have expectation 0, the non-0 element being invariant) and proposition 9 (which states a similar result for microbalanced random stimuli) provide access to a large family of empirically useful drift-balanced random stimuli. Instances that display striking apparent motion may be constructed readily.

In Section 8 we introduced microbalanced random stimuli, a distinguished subclass of drift-balanced random stimuli defined by the following property: A random stimulus I is

microbalanced iff, for any space-time-separable function W , the product WI is drift balanced. Thus I is guaranteed to avoid systematically stimulating any Fourier power motion mechanisms encountering I through any space-time-separable window. It was proved that (proposition 5) any space-time-separable random stimulus is microbalanced; that (proposition 6) any invariant microbalanced stimulus is space-time separable; that (proposition 7) the product of two independent microbalanced random stimuli is microbalanced; that (proposition 9) any linear combination of pairwise independent microbalanced random stimuli, all but at most one of which has expectation 0, is microbalanced; and that (proposition 10) the spatiotemporal convolution of two independent microbalanced random stimuli is microbalanced. An implication of proposition 9 is that all the demonstration stimuli presented in this paper are not only drift balanced but also microbalanced. Finally (in proposition 11), we showed that the expected response of any elaborated Reichardt detector to any microbalanced random stimulus is 0 at any instant in time.

In light of earlier observations,⁷⁻¹⁴ the existence of non-Fourier mechanisms is hardly surprising. Such mechanisms have, however, received no thorough investigation. The range of types of such mechanisms has not yet been elaborated, and their psychophysical properties remain largely unstudied. The importance of proposition 3 and the results of Section 8 lies in their utility for constructing stimuli for probing both the nature of non-Fourier motion-detection mechanisms as well as the interaction between such mechanisms and the band-tuned motion detectors that were the focus of most previous research.

APPENDIX A

In this appendix we verify that $E[|\bar{I}(\omega, \theta, \tau)|^2]$ exists for any random stimulus I and any $\omega, \theta, \tau \in \mathbb{R}$ (which was presumed in definition 2). Let $D = \{(x, y, t) \in \mathbb{Z}^3 \mid I[x, y, t] \neq 0\}$; then

$$\begin{aligned} E[|\bar{I}(\omega, \theta, \tau)|^2] &= \int_{\mathbb{R}^{|\mathcal{D}|}} \sum i[x, y, t]i[p, q, r] \\ &\quad \times \exp\{j(\omega(x-p) + \theta(y-q) + \tau(t-r))\}f(i)di \\ &= \sum \int_{\mathbb{R}^{|\mathcal{D}|}} i[x, y, t]i[p, q, r]f(i)di \\ &\quad \times \exp\{j(\omega(x-p) + \theta(y-q) + \tau(t-r))\}, \end{aligned}$$

where each sum ranges over all pairs of points, $(x, y, t), (p, q, r) \in \mathbb{Z}^3$. Note now that

$$\int_{\mathbb{R}^{|\mathcal{D}|}} i[x, y, t]i[p, q, r]f(i)di = E[I[x, y, t]I[p, q, r]].$$

However, as a consequence of the (probabilistic version of the) Schwartz inequality,³⁷ we note that

$$E[I[x, y, t]I[p, q, r]] \leq (E[I[x, y, t]^2]E[I[p, q, r]^2])^{1/2}.$$

However, by the definition of a random stimulus, the two expectations on the right-hand side of the inequality exist. Hence $E[|\bar{I}(\omega, \theta, \tau)|^2]$ exists for all $\omega, \theta, \tau \in \mathbb{R}$.

APPENDIX B

In this appendix we prove lemma 1, which is as follows:

Let S be a random stimulus equal to the sum of a set Ω of pairwise independent random stimuli; then

$$E[|\bar{S}|^2] = |\bar{E}_S|^2 + \sum_{I \in \Omega} E[|\bar{N}_I|^2],$$

where $N_I = I - E_I$ for each $I \in \Omega$.

First we write

$$S = \sum_{I \in \Omega} (E_I + N_I).$$

The linearity of Fourier transformation then yields

$$\bar{S} = \sum_{I \in \Omega} (\bar{E}_I + \bar{N}_I).$$

Thus

$$|\bar{S}|^2 = \sum (\bar{E}_I(\bar{E}_I)^* + \bar{N}_I(\bar{N}_I)^* + \bar{E}_I(\bar{N}_I)^* + \bar{N}_I(\bar{E}_I)^*),$$

where the sum is over all $I, J \in \Omega$.

Note first, however, that, whenever $I \neq J$,

$$E[\bar{E}_I(\bar{E}_J)^* + \bar{N}_I(\bar{N}_J)^* + \bar{E}_I(\bar{N}_J)^* + \bar{N}_I(\bar{E}_J)^*] = \bar{E}_I(\bar{E}_J)^*,$$

since I and J are independent and

$$E[\bar{N}_I] = E[(\bar{N}_I)^*] = \bar{0}.$$

Moreover, whenever $I = J$,

$$\begin{aligned} E[\bar{E}_I(\bar{E}_I)^* + \bar{N}_I(\bar{N}_I)^* + \bar{E}_I(\bar{N}_I)^* + \bar{N}_I(\bar{E}_I)^*] \\ = \bar{E}_I(\bar{E}_I)^* + E[\bar{N}_I(\bar{N}_I)^*]. \end{aligned}$$

Thus

$$\begin{aligned} E[|\bar{S}|^2] &= \sum_{J \in \Omega} \sum_{K \in \Omega} \bar{E}_J(\bar{E}_K)^* + \sum_{I \in \Omega} E[|\bar{N}_I|^2] \\ &= \left| \sum_{J \in \Omega} \bar{E}_J \right|^2 + \sum_{I \in \Omega} E[|\bar{N}_I|^2] \\ &= |\bar{E}_S|^2 + \sum_{I \in \Omega} E[|\bar{N}_I|^2]. \quad \blacksquare \end{aligned}$$

APPENDIX C

In this appendix we prove that the random stimuli G and H of demonstrations 5 and 4 are drift balanced. These random stimuli stem from proposition 3. To make the bridge explicit, we shall need to derive a corollary (C1) that depends on the following lemma.

Lemma C1

For $M \in \mathbf{Z}^+$, let the random variables $\rho_0, \rho_1, \dots, \rho_{M-1}$ be pairwise independent, each uniformly distributed on $[-\pi, \pi)$; then, for any $x, y, t \in \mathbf{Z}$, define the random stimulus I by setting

$$I[x, y, t] = \sum_{m=0}^{M-1} d_m[x, y] (\cos(\rho_m)h_m[t] - \sin(\rho_m)k_m[t]),$$

where, in each case, $d_m, h_m,$ and k_m are all real-valued functions that equal zero at all but a finite number of points of their respective domains. I is then drift balanced.

Proof

For $m = 0, 1, \dots, M-1$, term m of I is space-time separable and hence drift balanced. Moreover, for each m , the expectations of $\sin(\rho_m)$ and $\cos(\rho_m)$ are both 0. Thus the expectation of each term of the sum yielding I is 0; the result follows from proposition 3. \blacksquare

We apply lemma C1 to prove the following corollary used in constructing stimuli for demonstrations 4 and 5.

Corollary C1

For $M, N \in \mathbf{Z}^+$, let $\rho_0, \rho_1, \dots, \rho_{M-1}$ be pairwise independent random variables, each uniformly distributed on $[-\pi, \pi)$; then, for any $x, y, t \in \mathbf{Z}$, define the random stimulus I by setting

$$I[x, y, t] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} d_m[x, y] p_{m,n}[t] \cos(q_{m,n}[t] + \rho_m),$$

where, for $m = 0, 1, \dots, M-1$, and $n = 0, 1, \dots, N-1$, the functions $d_m, p_{m,n}$, and $q_{m,n}$ are real valued and zero at all but a finite number of points of their corresponding domains. I is then drift balanced.

Proof

We recast I so as to apply lemma C1:

$$\begin{aligned} I[x, y, t] &= \sum_{m=0}^{M-1} d_m[x, y] \sum_{n=0}^{N-1} p_{m,n}[t] \\ &\quad \times (\cos(q_{m,n}[t])\cos(\rho_m) - \sin(q_{m,n}[t])\sin(\rho_m)) \\ &= \sum_{m=0}^{M-1} d_m[x, y] (h_m[t]\cos(\rho_m) - k_m[t]\sin(\rho_m)) \end{aligned}$$

for

$$h_m[t] = \sum_{n=0}^{N-1} p_{m,n}[t] \cos(q_{m,n}[t]),$$

$$k_m[t] = \sum_{n=0}^{N-1} p_{m,n}[t] \sin(q_{m,n}[t]). \quad \blacksquare$$

Proof That H (Demonstration 4) Is Drift Balanced

H contains N frame blocks indexed $0, 1, \dots, N-1$, each composed of M rectangles indexed $0, 1, \dots, M-1$ from left to right. Let $\rho_0, \rho_1, \dots, \rho_{M-1}$ be pairwise independent random variables, each uniformly distributed on $[-\pi, \pi)$. Let C be a contrast value. We can express H as follows: For $m = 0, 1, \dots, M-1$, let $d_m[x, y] = 1$ for (x, y) in the m th rectangle and 0 elsewhere, and for $n = 0, 1, \dots, N-1$, let $g_n[t] = 1$ in the n th frame block and 0 elsewhere; then

$$H[x, y, t] = C \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} d_m[x, y] g_n[t] \times \cos\left(4\pi\left(1 + \cos\left(2\pi\left(\frac{m}{M} - \frac{n}{N}\right)\right)\right) + \rho_m\right).$$

To check that H is drift balanced, make the following identifications, and then apply corollary C1:

$$p_{m,n}[t] = C g_n[t]$$

and

$$q_{m,n}[t] = 4\pi\left(1 + \cos\left(2\pi\left(\frac{m}{M} - \frac{n}{N}\right)\right)\right).$$

Thus corollary C1 applies, and we conclude that H is drift balanced. (Note that H does not exploit the full generality of corollary C1, since, for these identifications, $p_{m,n}[t]$ does not depend on m and $q_{m,n}[t]$ does not depend on t .)

Proof That G (Demonstration 5) Is Drift Balanced

The random stimulus G is made up of N frame blocks indexed $0, 1, \dots, N-1$, each containing M rectangles indexed $0, 1, \dots, M-1$ from left to right. Let $\rho_0, \rho_1, \dots, \rho_{M-1}$ be pairwise independent random variables, each uniformly distributed on $[-\pi, \pi]$. Let C be some contrast value. We can then express G as follows: For $m = 0, 1, \dots, M-1$, let $d_m[x, y] = 1$ for (x, y) in the m th rectangle and 0 elsewhere; for $n = 0, 1, \dots, N-1$, let $g_n[t] = 1$ for t in the n th frame block and 0 elsewhere; then

$$G[x, y, t] = \frac{C}{2} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} d_m[x, y] g_n[t] \times \left(\cos\left(2\pi\left(\alpha \frac{m}{M} - \beta \frac{n}{N}\right)\right) + 1\right) \cos\left(2\pi\gamma \frac{n}{N} + \rho_m\right).$$

To see that G is drift balanced, set

$$p_{m,n}[t] = \frac{C}{2} g_n[t] \left(\cos\left(2\pi\left(\alpha \frac{m}{M} - \beta \frac{n}{N}\right)\right) + 1\right)$$

and

$$q_{m,n}[t] = 2\pi \frac{\gamma n}{N},$$

and apply corollary C1. (Note that, as with H_ρ , G_ρ does not exploit the full generality of corollary C1, since $q_{m,n}[t]$ depends on neither m nor t .)

ACKNOWLEDGMENTS

The research reported here was supported by U.S. Air Force Life Science Directorate, Visual Information Processing Program, grant 85-0364. The authors thank Michael S. Landy for helpful comments on various drafts of the manuscript.

REFERENCES

1. J. P. H. van Santen and G. Sperling, "Temporal covariance model of motion perception," *J. Opt. Soc. Am. A* 1, 451-473 (1984).

2. J. P. H. van Santen and G. Sperling, "Elaborated Reichardt detectors," *J. Opt. Soc. Am. A* 2, 300-321 (1985).
 3. E. H. Adelson and J. R. Bergen, "Spatiotemporal energy models for the perception of motion," *J. Opt. Soc. Am. A* 2, 284-299 (1985).
 4. A. B. Watson and A. J. Ahumada, "A look at motion in the frequency domain," NASA Tech. Memo. 84352 (National Aeronautics and Space Administration, Washington, D.C., 1983).
 5. D. J. Fleet and A. D. Jepson, "On the hierarchical construction of orientation and velocity selective filters," Tech. Rep. RBCV-TR-85-8 (Department of Computer Science, University of Toronto, Toronto, 1985).
 6. D. J. Heeger, "Model for the extraction of image flow," *J. Opt. Soc. Am. A* 4, 1455-1471 (1987).
 7. A. Pantle and L. Picciano, "A multistable movement display: evidence for two separate motion systems in human vision," *Science* 193, 500-502 (1976).
 8. M. Green, "What determines correspondence strength in apparent motion," *Vision Res.* 26, 599-607, 1986.
 9. A. M. Derrington and G. B. Henning, "Errors in direction-of-motion discrimination with complex stimuli," *Vision Res.* 27, 61-75 (1987).
 10. A. M. Derrington and D. R. Badcock, "Separate detectors for simple and complex grating patterns?" *Vision Res.* 25, 1869-1878 (1985).
 11. A. Pantle and K. Turano, "Direct comparisons of apparent motions produced with luminance, contrast-modulated (CM), and texture gratings," *Invest. Ophthalmol. Vis. Sci.* 27, 141 (1986).
 12. K. Turano and A. Pantle, "On the mechanism that encodes the movement of contrast variations. I. velocity discrimination," submitted to *Vision Res.*
 13. G. Sperling, "Movement perception in computer-driven visual displays," *Behav. Res. Methods Instrum.* 8, 144-151 (1976).
 14. J. T. Petersik, K. I. Hicks, and A. J. Pantle, "Apparent movement of successively generated subjective figures," *Perception* 7, 371-383 (1978).
 15. C. Chubb and G. Sperling, "Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception," *Invest. Ophthalmol. Vis. Sci.* 28, 233 (1987).
 16. The main demonstrations and results described herein were first reported at the Symposium on Computational Models in Vision, Center for Visual Science, University of Rochester, June 20, 1986, and at the meeting of the Association for Research in Vision and Ophthalmology, Sarasota, Florida, May 7, 1987.
 17. A. B. Watson and A. J. Ahumada, Jr., "Model of human visual-motion sensing," *J. Opt. Soc. Am. A* 2, 322-342 (1985).
 18. J. Victor, "Nonlinear processes in spatial vision: analysis with stochastic visual textures," *Invest. Ophthalmol. Vis. Sci.* 29, 118 (1988).
 19. C. Chubb and G. Sperling, "Texture quilts: basic tools for studying motion from texture," Publ. 88-1 of *Mathematical Studies in Perception and Cognition* (Department of Psychology, New York University, New York, 1988).
 20. W. Reichardt, "Autokorrelationsauswertung als Funktionsprinzip des Zentralnervensystems," *Z. Naturforschung Teil B* 12, 447-457 (1957).
 21. J. P. H. van Santen and G. Sperling, "Applications of a Reichardt-type model of two-frame motion," *Invest. Ophthalmol. Vis. Sci.* 25, 14 (1984).
 22. A. B. Watson, A. J. Ahumada, and J. E. Farrell, "The window of visibility: a psychophysical theory of fidelity in time-sampled motion displays," NASA Tech. Paper 2211 (National Aeronautics and Space Administration, Washington, D.C., 1983).
 23. A. B. Watson, A. J. Ahumada, Jr., and J. E. Farrell, "Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays," *J. Opt. Soc. Am. A* 3, 300-307 (1986).
 24. Grossberg S. and E. Mingolla, "Neural dynamics of form perception boundary completion, illusory figures and neon color spreading," *Psychol. Rev.* 92, 173-211 (1985).
 25. S. Grossberg and E. Mingolla, "Neural dynamics of form perception: textures, boundaries, and emergent segmentations," *Percept. Psychophys* 38, 141-171 (1985).
 26. R. J. Watt and M. J. Morgan, "The recognition and representation of edge blur: evidence of spatial primitives in human vision," *Vision Res.* 23, 1465-1477 (1983).

27. R. J. Watt and M. J. Morgan, "Spatial filters and the localization of luminance changes in human vision," *Vision Res.* **24**, 1387-1397 (1984).
28. R. J. Watt and M. J. Morgan, "A theory of the primitive spatial code in human vision," *Vision Res.* **25**, 1661-1674 (1985).
29. G. J. Burton, "Evidence of nonlinear response processes in the human visual system from measurements on the thresholds of spatial beat frequencies," *Vision Res.* **13**, 1211-1225 (1973).
30. A. Y. Maudarbocus and K. H. Ruddock, "Non-linearity of visual signals in relation to shape-sensitive adaptation responses," *Vision Res.* **13**, 1713-1737 (1973).
31. E. Peli, "Perception of high-pass filtered images," in *Visual Communications and Image Processing II*, T. R. Hsing, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **845**, 140-146 (1987).
32. T. Caelli, "Three processing characteristics of visual texture segmentation," *Spatial Vision* **1**, 19-30 (1985).
33. C. Chubb, G. Sperling, and D. H. Parish, "Designing psychophysical discriminations tasks for which ideal performance is computationally tractable," Publ. 88-2 of *Mathematical Studies in Perception and Cognition* (Department of Psychology, New York University, New York, 1988).
34. D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966), pp. 209-231.
35. C. Chubb and G. Sperling, "Processing stages in non-Fourier motion perception," *Invest. Ophthalmol. Vis. Sci.* **29**, 266 (1988).
36. G. A. Miller and W. G. Taylor, "The perception of repeated bursts of noise," *J. Acoust. Soc. Am.* **20**, 171-182 (1948).
37. W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1966), Vol. 2, p. 151.